UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Engenharia Elétrica e de Computação

Michelle Soares Pereira Facina

From 5G to 6G: Data Transmission Through Massive MIMO and Large Intelligent Surfaces

De 5G a 6G: Transmissão de Dados através de MIMO Massivo e Superfícies Inteligentes

Campinas

2020

Michelle Soares Pereira Facina

From 5G to 6G: Data Transmission Through Massive MIMO and Large Intelligent Surfaces

De 5G a 6G: Transmissão de Dados através de MIMO Massivo e Superfícies Inteligentes

Thesis presented to the School of Electrical and Computer Engineering of the University of Campinas in partial fulfillment of the requirements for the degree of Doctor in Electrical Engineering, in the area of Telecommunications and Telematics.

Tese apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Engenharia Elétrica, na Área de Telecomunicações e Telemática.

Supervisor: Prof. Dr. Gustavo Fraidenraich

Este exemplar corresponde à versão final da tese de doutorado da aluna Michelle Soares Pereira Facina, orientada pelo Prof. Dr. Gustavo Fraidenraich

---

Campinas

2020

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Área de Engenharia e Arquitetura
Rose Meire da Silva - CRB 8/5974

F118f

Facina, Michelle Soares Pereira, 1989-
From 5G to 6G : data transmission through massive MIMO and large intelligent surfaces / Michelle Soares Pereira Facina. – Campinas, SP : [s.n.], 2020.

Orientador: Gustavo Fraidenraich.
Tese (doutorado) – Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Sistemas de comunicação móvel. 2. Sistemas de comunicação móvel 5g. 3. Sistemas MIMO. 4. Materiais inteligentes. 5. Eficiência energética. I. Fraidenraich, Gustavo, 1975-. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: De 5G a 6G : transmissão de dados através de MIMO massivo e superfícies inteligentes
Palavras-chave em inglês:
Mobile communication systems
5g mobile communication systems
MIMO systems
Smart materials
Energy efficiency
Área de concentração: Telecomunicações e Telemática
Titulação: Doutora em Engenharia Elétrica
Banca examinadora:
Alvaro Augusto Machado de Medeiros
Claudio Ferreira Dias
Eduardo Rodrigues de Lima
Fabryccio Akkazza Chaves Machado Cardoso
Data de defesa: 01-12-2020
Programa de Pós-Graduação: Engenharia Elétrica

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0000-0002-0785-9606
- Currículo Lattes do autor: http://lattes.cnpq.br/3630752666583309

COMISSÃO JULGADORA - TESE DE DOUTORADO

Candidata: Michelle Soares Pereira Facina

Data da Defesa: 01 de dezembro de 2020

Título da Tese: "From 5G to 6G: Data Transmission Through Massive MIMO and Large Intelligent Surfaces

De 5G a 6G: Transmissão de Dados através de MIMO Massivo e Superfícies Inteligentes".

Prof. Dr. Gustavo Fraidenraich
Prof. Dr. Alvaro Augusto Machado de Medeiros
Dr. Cláudio Ferreira Dias
Dr. Eduardo Rodrigues de Lima
Dr. Fabryccio Akkazzha Chaves Machado Cardoso

A ata de defesa, com as respectivas assinaturas dos membros da Comissão Julgadora, encontra-se no SIGA (Sistema de Fluxo de Dissertação/Tese) e na Secretaria de Pós- Graduação da Faculdade de Engenharia Elétrica e de Computação.

Acknowledgements

This thesis work has been an intense learning experience for me. I was introduced to many new ideas for almost five years to improve mobile communication in the coming decades.

I am grateful to God for allowing me to overcome challenges and show me the strength I have.

I sincerely thank my supervisor, Dr. Gustavo Fraidenraich, for his guidance, support, and encouragement in many challenging situations. It took an incredible amount of time and effort to advise, discuss, suggest, and review each step. He worked hard not to let me give up, something that I will cherish for the rest of my career.

I am grateful to all my professors at UFJF and Unicamp, who gave me a wealth of knowledge and sparked my interest in wireless communication research.

To the professors Merouane Debbah and Marco di Renzo, thank you very much for the opportunity to study in France and expand my knowledge on intelligent surfaces.

My family members have been a constant source of support and joy in my life. Each of them contributed as best they could to get me here: my mother Sandra, my father Wellington, my sister Priscila, and especially my husband, always ready to resolve any difficulties that I may face, even remotely. His companionship and love are essential to me.

I am also grateful to friends from Wisstek and France, who filled the days with fun and laughter, making my Ph. D. a smooth experience.

Finally, I would like to thank the São Paulo Research Foundation (FAPESP) for the financial support provided by grants #2016/16181-2 and #2018/19538-4.

For my husband, Alex.

Abstract

The evolution of technology is fast. At the beginning of the doctoral project, the most discussed topic was the fifth generation of cellular networks. Now, in the end, we are already talking about the sixth generation and what comes next. With this in mind, my thesis addresses topics related to massive MIMO and Large Intelligent Surfaces (LIS). Despite the chronological distance between them (the first in 2016 and the second in 2020), both represent a significant leap in developing a low latency access network and a drastic change in the mobile communications infrastructure.

With numerous antennas on the base station and user, one of the problems inherent to massive MIMO is interference. Thus, we present an accurate approximation to the signal-to-interference ratio in two scenarios: when the number of antennas at the base station is finite or when this number is large enough to be considered infinite. We consider path loss and lognormal shadowing and prove the perfect match between simulation and our approximation from practical case studies.

Concerning LIS, we compare the system performance already published in the literature to that we obtained from an LoS propagation modeled of Rician channels and Matched Filter (MF) from preliminary analyses. We propose an accurate approximation for spectral efficiency and evaluate the path-loss when the base station does not have channel state information and transmits the symbols with equal powers to all users.

In the sequel, we do an in-depth analysis of a practical LIS-assisted Single-Input Single-Output (SISO) system in order to characterize its performance. We assume realistic imperfections for the values of reflection coefficients. In the first step, symbol error probability expressions are obtained for SISO scenarios with and without phase errors. We validate all the results by numerical simulations and have shown an excellent agreement. Comparisons with existing models show that practical systems present a considerable loss in performance.

In the next stage, the study is even more comprehensive. Since quantization errors are unavoidable, we evaluate the influence of the number of bits dedicated to the phase quantization on important performance such as spectral efficiency, symbol error rate, and outage probability. Based on Monte Carlo simulations, we prove our approach's excellent accuracy and investigate the behavior of the power scaling law and the power required to reach a specific capacity, depending on the number of reflecting elements. We show that a LIS with approximately fifty elements and four dedicated bits for phase quantization outperforms a conventional system's performance, i.e., a system without the assistance of a LIS.

As conclusion, we can say that the massive MIMO and LIS technologies offer substantial improvements and came to stay and coexist.

Keywords: 5G, 6G, massive MIMO, Large Intelligent Surfaces, channel performance.

Resumo

A evolução da tecnologia é rápida. No início do projeto de doutorado, o tema mais discutido era a quinta geração de redes celulares. Agora, no final, já estamos falando sobre a sexta geração e o que vem a seguir. Com isso em mente, minha tese aborda tópicos relacionados a MIMO massivo e *Large Intelligent Surfaces* (LIS). Apesar da distância cronológica entre eles (o primeiro em 2016 e o segundo em 2020), ambos representam um salto significativo no desenvolvimento de uma rede de acesso de baixa latência e uma mudança drástica na infraestrutura de comunicações móveis.

Com várias antenas na estação base e no usuário, um dos problemas inerentes ao MIMO massivo é a interferência. Assim, apresentamos uma aproximação precisa da razão sinal-interferência em dois cenários: quando o número de antenas na estação base é finito e quando esse número é grande o suficiente para ser considerado infinito. Consideramos a perda de percurso e o sombreamento lognormal e provamos o excelente casamento entre simulação e nossa aproximação.

Em relação a LIS, comparamos o desempenho do sistema já publicado na literatura com o que obtivemos de uma propagação baseada em linha de visada modelada a partir de canais Rice e *Matched Filter* (MF). Propomos uma aproximação precisa para eficiência espectral e avaliamos a perda quando a estação base não possui nenhuma informação de estado do canal e transmite os símbolos com potências iguais para todos os usuários.

Em seguida, fazemos uma análise aprofundada de um sistema prático Single-Input Single-Output (SISO) assistido por LIS de forma a caracterizar seu desempenho. Assumimos imperfeições para os valores dos coeficientes de reflexão. Em uma primeira etapa, as expressões de probabilidade de erro de símbolo são obtidas para cenários SISO com e sem erros de fase. Validamos todos os resultados por simulações numéricas e demonstramos uma excelente concordância. As comparações com os modelos existentes mostram que os sistemas práticos apresentam uma perda considerável de desempenho.

Na próxima etapa, o estudo é ainda mais abrangente. Como os erros de quantização são inevitáveis, avaliamos a influência do número de bits dedicados à quantização de fase no cálculo de importantes parâmetros de desempenho, tais como, eficiência espectral, taxa de erro de símbolo e probabilidade de interrupção. Com base em simulações de Monte Carlo, provamos a excelente precisão de nossa abordagem e investigamos o comportamento de power scaling-law e a potência necessária para atingir uma capacidade específica, dependendo do número de elementos reflectores. Mostramos que um LIS com aproximadamente cinquenta elementos e quatro bits dedicados à quantização de fase supera o desempenho de um sistema convencional, ou seja, um sistema sem o auxílio de LIS.

Como conclusão, podemos dizer que as tecnologias MIMO massivo e LIS oferecem melhorias substanciais e vieram para ficar e coexistir.

Palavras-chaves: 5G, 6G, MIMO massivo, superfícies inteligentes, desempenho do canal.

List of Figures

List of Tables

List of Abbreviations

5G	fifth generation
6G	sixth generation
AWGN	additive white Gaussian noise
BER	bit error rate
BPSK	binary phase shift keying
BS	base station
CDF	cumulative distribution function
CLT	central limit theorem
CSI	channel state information
EE	energy efficiency
IoE	internet of everything
IoT	internet of things
LIS	large intelligent surfaces
LoS	line-of-sight
LTE	long-term evolution
MF	matched filter
MRC	maximal ratio combining
MIMO	multiple-input multiple-output
MM	majorization maximization
mmWave	milimeter wave
NOMA	non-orthogonal multiple access
OFDM	orthogonal frequency division multiplexing
PDF	probability density function
RF	radio frequency
RIS	reconfigurable intelligent surfaces
SE	spectral efficiency
SER	symbol error rate
SINR	signal-to-interference-plus-noise ratio
SIR	signal-to-interference ratio
SISO	single-input single-output
SNR	signal-to-noise ratio
TDD	time division duplex
WiMax	worldwide interoperability for microwave access
ZF	zero-forcing

List of Symbols

α_i	phase of link between BS and the i-th element of LIS
$\bar{a}$	constant modulation dependent parameter
a_c	parameter that controls how strongly the users cluster towards the BS
a_e	parameter that controls how strongly the users cluster towards the edge cell
A_1, A_2	auxiliary parameters
B	total bandwidth in Hz
b	number of quantization bits
$\bar{b}$	constant modulation dependent parameter
β_{jkl}	long-term fading coefficient
β_{θ_i}	amplitude on the combined incident signal
β_1	large scale fading between BS and LIS
β_g	large scale fading between BS and LIS
β_2	large scale fading between LIS and user
β_h	large scale fading between LIS and user
$\beta_{\mathbf{min}}$	minimum amplitude of LIS
C	sum-capacity in Chapter 3
$\bar{C}$	average sum-capacity
$\bar{C}_{\text{inst.}}$	instantaneous spectral efficiency
$\bar{C}_{\text{upper}}$	upper bound for ergodic spectral efficiency
$\bar{C}_{\text{upper}}^{\text{max.}}$	maximum upper bound for ergodic spectral efficiency
$\bar{C}_{\text{lower}}$	lower bound for ergodic spectral efficiency
$\bar{C}_{\text{lower}}^{\text{max.}}$	maximum lower bound for ergodic spectral efficiency
$\bar{C}_{\text{high-SNR}}$	ergodic spectral efficiency in high SNR regime
$\bar{C}_{\text{high-SNR,N}}$	ergodic spectral efficiency in high SNR and N regimes
C_{jk}^{u}	net capacity per terminal for uplink
d_{jkl}	distance between k-user of l-cell and the BS in j-cell
D	diversity order
δ_i	phase of link between the i-th element of LIS and user
Δ	reuse frequency pattern
Δf	frequency bandwidth
ϵ	decay exponent
ϵ	spectral efficiency degradation in bits/s/Hz in Chapter 5
E_s	average transmitted energy per symbol
$f_{d_{jkl}}(d)$	probability density function of d_{jkl}
$f_{C_{inst.}}(d)$	probability density function of $_{inst.}$
$f_X(x)$	probability density function of x

$f_\gamma(\gamma)$	probability density function of γ
$F_\gamma(\gamma)$	cumulative density function of γ
$F_{C_{inst.}}$	cumulative density function of $C_{inst.}$
g_k	precoding vector
$\tilde{g}_{jkl}$	short-term fading coefficient
γ	transmit power loss due to cyclic prefix
γ_{jk}	uplink signal-to-interference ratio at j-th BS of k-user
$\boldsymbol{H_1}$	channel gain matrix between BS and LIS
g_n	channel gain coefficient between BS and n-th element of LIS
$\boldsymbol{H}$	channel response in Chapter 4
$\boldsymbol{H_{i1}}$	channel gain vector between BS and i-th element of LIS
h_n	channel gain coefficient between n-th element of LIS and user
$\boldsymbol{H_2}$	channel gain matrix between LIS and users
$\boldsymbol{H_{i2}}$	channel gain vector between i-element of LIS and user
$\boldsymbol{h_{2,k}}$	channel gain vector between LIS and k-user
k	parameter that controls the steepness of the function curve
K	number of user terminals
κ	shape parameter of Gamma distribution in Chapter 5
K_F	shape parameter
L	number of BSs
M	number of BS antennas
μ	mean of βjkl in Chapter 2
$_y$	approximated mean of y in Chapter 2
$_z$	approximated mean of z in Chapter 2
$_\beta$	mean of β in Chapter 4
$_H$	mean of H in Chapter 4
N	number of BSs in Chapter 2 and number of LIS elements in Chapter 3
n	number of LIS elements from Chapter 4
N_j	AWGN matrix
n_j	element of the AWGN matrix
ν	path-loss exponent
p_k	transmit power
ϕ	horizontal distance between $\frac{-\pi}{2}$ and $\beta_{\min}$
ϕ_n	adjustable phase-shift produced by the n-th LIS reflector
Φ	diagonal matrix of phase-shift produced by the n-th LIS reflector
P_{BS}	total hardware static power comsumption at BS
P	transmit power
Ψ_k	assigned sequence for the k-th user
P_{max}	maximal system power

P_n	total hardware static power consumption at LIS
P_{UE}	total hardware static power dissipated by user
P_e	symbol error probability
Pr	probability of a given event occurs
$P_{out.}$	outage probability
$P_{out.}^{\text{high-SNR}}$	outage probability in high SNR regime
P_e	symbol error probability
$P_e^{\text{high-SNR}}$	symbol error probability in high SNR regime
ρ_p	uplink pilot transmit power
ρ_u	uplink transmit power
ρ	average SNR
$R_{min,k}$	individual QoS of the k-user
r_c	radius of a hexagonal cell
r_g	distance between source and LIS
r_h	distance between LIS and destination
r	inner radius
R	outer radius
s	modulation data symbol
s_k	unit power complex-valued information symbol
σ_n^2	noise variance in Chapter 2
$\overset{2}{}_x$	approximated variance of x in Chapter 2
$\overset{2}{}_y$	approximated variance of y in Chapter 2
$\overset{2}{}_z$	approximated variance of z in Chapter 2
σ^2	variance of βjkl in Chapter 2 and variance of $\boldsymbol{H_1}$ and $\boldsymbol{H_2}$ in Chapter 3
σ_w^2	variance of noise
σ_W^2	variance of noise in Chapter 4
σ_ϕ^2	variance of ϕ
σ_β^2	variance of β in Chapter 4
σ_H^2	variance of H in Chapter 4
$\sigma_{\mathbf{shadow}}$	standard deviation of z_g and z_h in Chapter 5
τ	available pilot sequences
θ_n	effective phase shifts applied by n-th LIS reflecting element
θ_i	effective phase shifts applied by i-th LIS reflecting element
θ	scale parameter of Gamma distribution in Chapter 5
$\mathbf{T_{pilot}}$	time to transmit pilot sequences
$\mathbf{T_{slot}}$	slot length
$\mathbf{T_u}$	useful symbol duration
x	binary transmitted signal
x_{kl}	data symbol from the k-user of the l-cell

ξ	amplifier efficiency
Y_j^p	matrix of received signal at j-th BS
y_j^p	received signal at j-th BS
y_j^p	received signal in Chapter 4
y	received signal in Chapter 5
w	additive white Gaussian noise
w_k	thermal noise
z_g	random variable of link between source and LIS
z_h	random variable of link between LIS and destination

Contents

1

Introduction

This doctorate was born in 2016 when the fifth generation of mobile communications (5G) was still a distant reality. The text of the initial project written at that time quoted the following:

"Technological advances such as real-time multimedia, video, and high-resolution image applications have driven demand for ever-increasing communications service fees. A growth factor of approximately forty times in the next five years is estimated [4]. However, this demand arises, concomitantly, in a scenario where the availability of frequency and energy spectrum is increasingly limited [5].

To supply the growing demand, new researches have been made to include fundamental changes for 5G [6]. The next generation is expected to bring some changes from the current generation. The most important technical improvements aim to achieve the following aspects: higher data transmission speed, lower latency, and greater energy efficiency. The main idea is to guarantee the necessary structure for the Internet of Things (IoT) to become a reality in the world [7].

Thus, the big question still unanswered is how to offer increased throughput over the wireless network reliably and uniformly. The main solutions already presented in the literature cover three categories:

- Exploration of the unused frequency spectrum;
- Implementation of more access points, specific to more restricted areas;
- Use of access points and terminals with multiple antennas.

Specifically, our goal is to study the use of multiple antennas added to both the Base Station (BS) and mobile terminal. Massive Multiple-Input Multiple-Output (MIMO), also known as *Large-Scale Antenna Systems*, *Large-Scale MIMO*, ARGOS, *Full-Dimension* MIMO and *Hyper*-MIMO, is an emerging technology scaled to orders of magnitude far superior to the current technical scenario."

At that time, one of the news that motivated me to study Massive MIMO was that researchers at Rice University (Texas) conducted experimental research on even more comprehensive wireless technology. Known as ArgosNet, the prototype that already exists (Figure

1.1) is equipped with up to twelve programmable base stations, each with about 100 or more antennas [1].

Figure 1.1 – Argos-Net prototype [1].

But the evolution of technology is fast. Now, in 2020 5G represents a significant leap in developing a low latency access network through new frequency bands such as the millimeter (mmWave) wave spectrum, licensed and unlicensed bands [8]. If it became a commercial success, it was thanks to massive MIMO.

Digitally controllable antenna arrays can be deployed "anywhere." Conventional sites operating in the sub-6 GHz band are equipped with arrays of at least 64 antennas (per sector) for spatial multiplexing over broad areas. New BSs operating on mmWave bands can be deployed both indoor and at street level to provide coverage of the local area [9]. In fact, according to Zappone et al. [10], our society is undergoing a digitization revolution, with a substantial increase in internet users and connected devices. As innovative as the fifth generation, features such as infrastructure and antenna densification, and frequency bands in the mmWave range system cannot achieve the desired requirements. Its implantation is continually exposing some limitations, and the original premise as an enabler for the Internet of Everything (IoE) is not possible [11].

To overcome these challenges, a disruptive sixth generation (6G) wireless, whose design follows technological trends and is adapted to the performance requirements of IoE applications is being considered. Despite recent initiatives, the architecture and performance components of 6G remain mostly undefined.

A possible candidate to join such a generation is a Large Intelligent Surface (LIS). It is a planar array comprising many nearly passive, low-cost, reflecting elements with reconfigurable parameters. Figure 1.2, extracted from [2], provides a more practical overview of implementation. It is composed of some unit cells to form an array surface. Figure 1.2(a) illustrates the unit cell's top view schematic and its simplified equivalent circuit model. Four metallic rectangular patches, each pair of which is bridged by a varactor diode, constitute a unit cell. The varactor

diodes are biased via holes by the feeding network, composed of a slotted copper plate at the substrate's bottom. The unit cell can be modeled as a parallel resonant tank. The C, R, L1, L2, Zl, Z0 and represent the equivalent capacitance, resistance, inductance on the top of the unit cell, the equivalent inductance at the bottom of the unit cell, the equivalent load impedance of the unit cell, the characteristic impedance of the air and the reflection coefficient of the unit cell, respectively. The capacitance of the simplified equivalent circuit model for the unit cell is dominated by the varactor diode, which indicates that the load impedance can be tuned by the biasing voltage of the varactor diode.

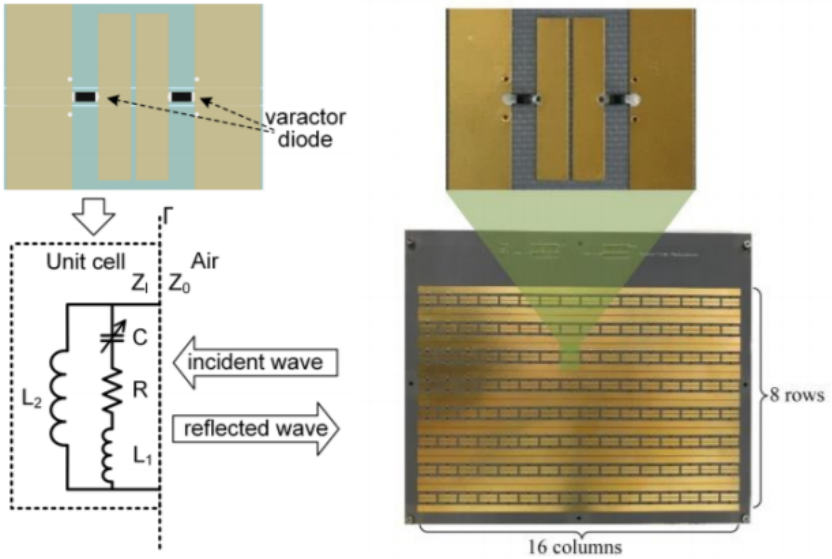


Figure 1.2 – An overview of the LIS implementation [2].

With the smart controller's help, it assists in interconnecting the physical and digital worlds seamlessly and sustainably. This way, the wireless environment itself is turned into a software-reconfigurable entity.

The uncertainties are countless, but we can guarantee that 6G will confluence past trends and emerging trends with several technologies coexisting. So, it is only fair to study each one of them. There is an envision that the LIS technology plays a critical role in 6G communication networks, similar to the massive MIMO in 5G networks. Then, at least, let's study the two most important. So, this thesis addresses specific modeling topics about Massive MIMO and LIS. In the sequel, this chapter gives an overview of the state of the art of both technologies.

1.1 Massive MIMO

In its original design, known as point-to-point MIMO, the transmitter transmits data to a receiver equipped with some antennas. The users are accommodated at different times or

frequency blocks. Each use of the channel comprises sending and receiving a signal vector. Each signal received is a linear combination of transmitted signals.

Regarding point-to-point MIMO communication, there are already designed techniques capable of achieving diversity gains that improve the channel's reliability and transmission rate. Implementations of these techniques include time-space codes, antenna selection, and spatial multiplexing, which maximizes the data rate, and even capable techniques combining diversity and multiplexing gains [12].

On the other hand, multiuser MIMO has been extensively studied during the last two decades and applied to several wireless standards. It can significantly improve the system's performance. Typically, a BS equipped with some antennas serves users simultaneously, sharing the multiplexing gain. In addition to the equipment user terminals are cheaper, due to multiplexing, the system generally becomes less sensitive to the environment.

Within this context, transmitting and receiving antennas is one of the main attributes of modern wireless communication standards. For example, Long Term Evolution (LTE), Wi-Fi, and Worldwide Interoperability for Microwave Access (WiMAX) can achieve high energy and spectral efficiency. However, for most MIMO implementations, BS typically employs few antennas, and the corresponding improvement in spectral efficiency, while important, is still relatively modest.

The most exciting scenario in recent years has been the massive multiuser MIMO. In the last effort to achieve even more significant gains and simplify the required signal processing, a BS, equipped with a considerable number of antennas, serves a certain number of mobile terminals.

However, some aspects limit the performance of massive MIMO systems and deserve to be highlighted. Among them, we can mention the acquisition of Channel State Information (CSI), the detection and precoding algorithms, the fading channel, and the pilot contamination. We will focus our attention on the last two, detailing them further.

1.2 Large Intelligent Surfaces

The numerous advantages of massive MIMO [7], [13] are undeniable. So much so that it is already one of the requirements set out in the recently defined 5G standard [14]. However, the significant disadvantage of all this concerns the implementation challenges due to the enormous physical size of antenna arrays [15]. The propagation at high frequencies through millimeter waves has drastically reduced the size of the antennas. Still, they require many Radio Frequency (RF) components, limiting developments in the commercial application. Therefore, it is necessary to rethink antenna arrays' architecture to design solutions adapted explicitly to

more realistic scenarios.

LIS can be seen as a more natural evolution of massive MIMO but taken to the extreme. Because it is a surface with radiating and sensory elements, these surfaces resemble wallpapers electromagnetically active with built-in processing power. Its great advantage is to make it possible to focus energy in three-dimensional space and remote sensing with extreme precision.

A first analysis of the information transfer resources has already been carried out and proves this new tool's viability. However, little is known about LIS, and some of its parameters need to be correctly estimated.

1.3 Summary of Contributions and Thesis Outline

In order to present the evolution, this thesis is organized in chronological order of development and follows the structure below.

Chapter 2: From closed-form expressions, we model and analyze the performance of massive MIMO systems from the SIR point of view. Unlike most works that consider users uniformly distributed, we also take into account specific cases where they are concentrated in the center or the edge of the cell.

Chapter 3: Preliminary studies on optimization help us to understand how RIS-assisted systems work. They gave us the background for a more in-depth approach.

Chapter 4: Faced with real scenarios with channel estimation and quantization errors, we model some performance parameters of RIS-assisted systems.

Chapter 5: With further analysis, we again derived performance parameters from RIS-assisted systems, but now focusing our attention only on quantization errors.

2

Accurate Lognormal Approximation to the Signal-to-Interference Ratio in Massive MIMO

2.1 Introduction

MIMO emerges as a promising technology to meet the stringent requirements of future fifth generation systems. This technology represents a drastic change in the infrastructure of cellular networks, since its communication system is composed of Base BS with hundreds of antennas and can serve simultaneously dozens of user terminals, each having a single antenna. The BS is responsible for sending independent data streams to multiple user terminals in the same time-frequency resource [13]. Furthermore, each user terminal is ideally assigned with an orthogonal pilot sequence in the uplink channel during the training stage. However, the maximum number of such sequences is limited by the duration of the coherence interval. So, the available amount of orthogonal pilot sequences, in a multicell system, is finite and can result in pilot contamination in two different situations.

In the first, when the number of pilot sequences is superior or equal to the number of user terminals, it is assumed orthogonality between pilot sequences of the same cell. But the frequency is reused and employed in a regular pattern to ensure that the intercell interference remains below a harmful level. On the other hand, when the number of pilot sequences is inferior to the number of user terminals, pilot sequences are reused within the same cell to reduce the training overhead [16]. In this way, non-orthogonal pilot sequences need to be employed, which is the major source of pilot contamination, known as intracell interference [13].

The influence of the all the interference is normally modeled in the Signal-to-Interference Ratio (SIR) parameter. As our main contribution, this work presents an accurate log-normal approximation to the Cumulative Distribution Function (CDF) of SIR in two scenarios: when the number of antennas at BS is finite and when this number is large enough and can be considered infinite. Moreover, an approximation for the capacity has also been presented, in which the slow fading is a combination of path loss and log-normal shadowing. Practical case studies are provided to illustrate the very good match between simulation results and our analytical approximation. The validity of our approximation is extended not only to uniform

user distribution but also to users distributed non-uniformly. To the best of our knowledge, no similar results have been found in the literature.

The remaining of this paper is organized as follows: Section 2.2 presents the system model while in Section 2.3 closed-form expressions for SIR and system capacity are proposed. Section 2.4 shows some examples of a common scheme in the literature and final remarks are given in Section 2.5.

2.2 System Model

Based on Marzetta's work [7], our model, as shown in Figure 2.1, consists of a hexagonal cellular geometry. The BSs, located at the center of the cell, hold a determined number of antennas, M. For operation simplicity, the Orthogonal Frequency Division Multiplexing (OFDM)/Time Division Duplex (TDD) is used. In this way, it is assumed that the channel is flat fading and exhibits a reciprocity behavior, reducing the overhead required for the acquisition of Channel State Information (CSI) by means of uplink training signals. The OFDM symbol interval is denoted by T_{s}, the subcarrier spacing by Δf, and the useful symbol duration by $T_{\mathrm{u}} = 1/\Delta f$. The scheme consists of a tessellation of non cooperative hexagonal cells. The N

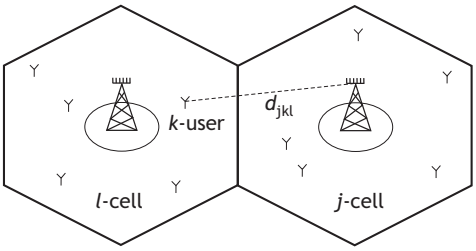


Figure 2.1 – System model with hexagonal cells and, the distance d_{jkl} between k-user of l-cell and the BS in j-cell.

BSs are deterministically distributed in a circular region. It is assumed that the BSs employ frequency reuse Δ and that the BSs in different bands do not interfere with each other. There are, in total, K user terminals within each of the L active cells re-using the same frequency bands and pilot sequences. The system performance is studied under the identically and independently distributed (iid) fading assumption, including uncorrelated log-normal shadow fading, and geometric attenuation.Also, the channels are considered spatially uncorrelated with each other. The user terminals are randomly distributed in a hexagonal cell, whose radius is r_c, except in a disk of radius 100 m centered at the BS. Interfering nodes, that affect a particular cell, are separated into tiers that reuse the same frequency band and are within eight cell-diameters of that cell, as adopted by Marzetta [7]. Also, it is assumed perfect synchronism between the signals received from different cells, which is the worst case in terms of pilot contamination.

As in [17], for each subcarrier, the vector between the j-th BS and the k-th user at the l-th cell is denoted by $g_{jkl} = \sqrt{\beta_{jkl}}\tilde{g}_{jkl}$, in which β_{jkl} refers to the long-term fading coefficient,

comprising path loss and log-normal shadowing. $\tilde{g}_{jkl}$ is the short-term fading channel vector, that follows a normal distribution with zero mean and unitary variance. In its turn, the orthogonal pilot sequences set is represented by $\Psi = [\Psi_1 \Psi_2 \dots \Psi_\tau] \in \mathbb{C}^{\tau \times \tau}$, in which τ is the number of available sequences, $\Psi^H \Psi = I_\tau$ and ^{H} is hermitian operator. Considering Ψ_k as the assigned sequence for the k-th user, the received signal at the j-th BS during the training stage is

$$Y_j^p = \sqrt{\rho_p} \sum_{l=1}^{L} \sum_{k=1}^{\tau} g_{jkl} \Psi_k^H + N_j, \tag{2.1}$$

in which ρ_p is the uplink pilot transmit power, $N_j \in \mathbb{C}^{M \times \tau}$ is the Additive White Gaussian Noise (AWGN) matrix with iid elements following a complex normal distribution with zero mean and variance σ_n^2. The j-th BS estimates the k-th user CSI by correlating Y_j^p with Ψ_k. By acquiring such estimates, the BS is able to perform linear detection in uplink employing the Maximal Ratio Combining (MRC) scheme in frequency domain. So, the j-th BS receives the following signal

$$y_j = \sqrt{\rho_u} \sum_{l=1}^{L} \sum_{k=1}^{K} g_{jkl} x_{kl} + n_j, \tag{2.2}$$

in which ρ_u is the uplink data transmit power, x_{kl} is the data symbol from the k-th user of the l-th cell and n_j is the $M \times 1$ AWGN sample vector.

2.3 Proposed Approximation

Next, this work presents the assumptions and steps to derive our proposed approximation for the distribution of SIR, considering a finite and infinite number of antennas M. Without loss of generality, only the uplink SIR is considered, but the downlink case is very similar.

2.3.1 Finite Number of BS Antennas

From the estimates available at the j-th BS, and the received signal during the uplink data transmission stage (2.2), the uplink SIR, employing MRC, can be given by [17, Eq. 14],

$$\gamma_{jk}^{\mathrm{u}} = \frac{\zeta_{jkj}^2}{\sum_{l=1, l \neq j}^{L} \zeta_{jkl}^2 + \frac{\alpha_{jk}^2}{M} \left(\sum_{l=1}^{L} \sum_{i=1}^{K} \zeta_{jil} + \frac{\sigma_n^2}{\gamma \rho_u} \right)}, \tag{2.3}$$

in which the index jkl refers to a quantity related to the k-th terminal in the l-th cell and the BS in the j-th cell. $\zeta_{jkj} = \frac{\beta_{jkj}}{d_{jkj}^{\epsilon}}$, in which the variable β_{jkl} represents the shadow fading coefficient, modeled as a log-normal random variable with mean μ, and variance σ^2 in the logarithmic scale, d_{jkl} is the distance between the user terminal and the correspondent BS. The variable ϵ is the decay exponent, typically $\epsilon \geq 2$. The AWGN follows a complex normal distribution with zero mean and variance σ_n^2, $\alpha_{jk}^2 = \sum_{l=1}^{L} \zeta_{jkl} + \frac{\sigma_n^2}{\rho^p}$, and $\rho_u = 1$, since it is assumed that all users have unit transmit power. Also, γ accounts the transmit power loss due to Cyclic Prefix (CP).

Given a certain number of users inside the cell, it is assumed that all the users are uniformly distributed from the inner radius r to the outer radius R, for a given tier. Since the circumference of a circle is proportional to its radius, the Probability Density Function (PDF) $f_{d_{jkl}}(d)$ is also proportional to its radius, that is, $f_{d_{jkl}}(d) = ad$ for some constant a. The condition $\int_r^R f_{d_{jkl}}(d) = 1$ implies in $a = \frac{2}{R^2 - r^2}$, therefore

$$f_{d_{jkl}}(d) = \begin{cases} \frac{2d}{R^2 - r^2}, & \text{if } r \leq d \leq R \\ 0, & \text{otherwise.} \end{cases} \tag{2.4}$$

The computation of the exact distribution of (2.3) is very intricate, and possibly, the final solution would be useless due to its complexity. In order to corroborate this assertion, the exact distribution of the numerator of (2.3), defined as $\Theta = \beta_{jkj}^2 d_{jkj}^{-2\epsilon}$, is derived. As said before, the random variable β_{jkj} is log-normal distributed and the distribution of d_{jkj} is given by (2.4). Finally, the PDF of Θ can be computed as the ratio of two random variables [18], and given in an exact manner as

$$f_\Theta(\theta) = \frac{\theta^{-\frac{\epsilon+1}{\epsilon}} e^{\frac{2\left(\sigma^2 + \mu\epsilon\right)}{\epsilon^2}} \left[\operatorname{erf}\left(\frac{-\frac{1}{2}\epsilon(2\epsilon\log(R) + \log(\theta)) + 2\sigma^2 + \mu\epsilon}{\sqrt{2}\sigma\epsilon} \right) - \operatorname{erf}\left(\frac{-\frac{1}{2}\epsilon(2\epsilon\log(r) + \log(\theta)) + 2\sigma^2 + \mu\epsilon}{\sqrt{2}\sigma\epsilon} \right) \right]}{2\epsilon(r^2 - R^2)}, \tag{2.5}$$

where $\operatorname{erf}(\cdot)$ is the error function.

Unfortunately, this PDF does not have a closed form solution for its characteristic function, therefore, the computation of the PDF of the sum, presented in the denominator, would not be possible. In order to circumvent this problem, it is proposed to approximate Θ by a log-normal distribution X. Through the moment matching method, the mean $\hat{\mu}_x$, and the variance $\hat{\sigma}_x^2$ of X are estimated using the equalities $\mathbf{E}[\log X] = \mathbf{E}[\log \Theta]$ and also $\mathbf{V}[\log X] = \mathbf{V}[\log \Theta]$, where $\mathbf{E}[\cdot]$, and $\mathbf{V}[\cdot]$ are the expectation, and variance operator, respectively. Therefore, the following can be written

$$\begin{aligned} \hat{\mu}_x &= \mathbf{E}[\log(\Theta)] = 2\mathbf{E}[\log(\beta)] - 2\epsilon\mathbf{E}[\log(d_{jkl})] = \\ &= 2\mu - \frac{\epsilon\left\{r^2\left[1 - 2\log(r)\right] + R^2\left[1 - 2\log(R)\right]\right\}}{r^2 - R^2}, \end{aligned} \tag{2.6}$$

where $\mathbf{E}[\log(d_{jkl})]$ is evaluated with respect to the distribution of d_{jkl}, as defined in (2.4). In the same way, the variance $\hat{\sigma}_x^2$ can be computed as,

$$\begin{aligned} \hat{\sigma}_x^2 &= \mathbf{V}[\log(\Theta)] = 4\mathbf{V}[\log(\beta)] + 4\epsilon^2\mathbf{V}[\log(d_{jkl})] = \\ &= 4\sigma^2 + \frac{\epsilon^2\left\{r^4 - r^2R^2\left[2 + 4\log^2\left(\frac{R}{r}\right)\right] + R^4\right\}}{\left(r^2 - R^2\right)^2}. \end{aligned} \tag{2.7}$$

Then, as supposed, the PDF of X is the very well known log-normal distribution given as [18]

$$f_X(x) = \frac{e^{-\frac{(\log(x) - \hat{\mu}_x)^2}{2\hat{\sigma}_x^2}}}{\sqrt{2\pi}x\hat{\sigma}_x}. \tag{2.8}$$

The accuracy of the proposed approximation is illustrated in Figures 2.2 and 2.3. These figures show the approximate, simulated and exact CDF and PDF of the random variable Θ, respectively. As can be noted, our approximation is excellent and this is the key assumption for approximating the distribution of (2.3) as a log-normal distribution.

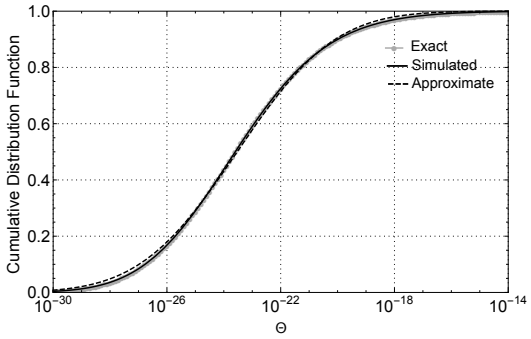


Figure 2.2 – Simulated, approximate and exact cumulative distribution function for Θ ($r = 100$ and $R = 2000$).

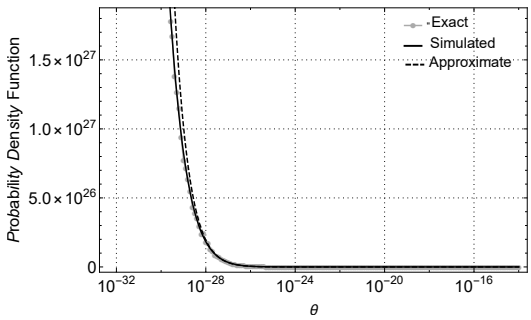


Figure 2.3 – Simulated, approximate and exact probability density function for Θ ($r = 100$ and $R = 2000$).

In the literature [19], it is well known that the ratio of two log-normal variables is also a log-normal variable. In this sense, if a log-normal distribution could approximate the distribution of the denominator of (2.3), then the entire ratio, given in (2.3), also would be log-normal distributed. To this end, every term inside the sum of the denominator is approximated as a log-normal distribution (similar to the X variable), and then, based on the result of Filho et al. [20], to approximate the sum of log-normal variables as a log-normal variable. This can be performed by matching the first two moments of the inverse exact sum with those of the inverse log-normal approximation. In this way, the parameters of the resulting log-normal distribution are determined.

Lets call the first term in the denominator of (2.3) as S. This random variable can be approximated as the sum of approximate iid log-normal variables as $S = \sum_{l \neq j} \beta_{jkl}^2 d_{jkl}^{-2\epsilon}$. We want to approximate the sum of log-normal variates in S by a single log-normal variable, denoted here as Y. To this end, the moments are matched in the following way: $\mathbf{E}[Y^{-1}] = \mathbf{E}[S^{-1}]$, and also $\mathbf{E}[Y^{-2}] = \mathbf{E}[S^{-2}]$. Since $\mathbf{E}[Y^{-1}] = e^{-\left(\hat{\mu}_y - \hat{\sigma}_y^2/2\right)}$ and $\mathbf{E}[Y^{-2}] = e^{-2\left(\hat{\mu}_y - 2\hat{\sigma}_y^2/2\right)}$, then with some

mathematical manipulations, it is possible to write that

$$\begin{aligned} \hat{\mu}_y &= 0.5 \ln \mathbf{E}[S^{-2}] - 2 \ln \mathbf{E}[S^{-1}] \\ \hat{\sigma}_y^2 &= \ln \mathbf{E}[S^{-2}] - 2 \ln \mathbf{E}[S^{-1}], \end{aligned} \tag{2.9}$$

where $\mathbf{E}[S^n]$ is the n-th moment of S, while $\hat{\mu}_y$ and $\hat{\sigma}_y$ are the mean and standard deviation of the approximated log-normal random variable Y, respectively.

Following the same reasoning, the second term in the denominator of (2.3) presents the terms $V = \frac{\alpha_{jk}^2}{M}$, and $W = \sum_{l=1}^{L} \sum_{k=1}^{K} \frac{\beta_{jkl}}{d_{jkl}^{\epsilon}} + \sigma_n^2$. Both terms are approximated as log-normal distributed, and since the product of log-normal distributions is also log-normal distributed, the random variable $Z = VW$ can be approximated as a log-normal variable. So, $\hat{\mu}_z = \mu_v + \mu_w$ and $\hat{\sigma}_z^2 = \sigma_v^2 + \sigma_w^2$, where μ_v, μ_w, σ_v^2 and σ_w^2 are means and variances of the variables V and W, respectively. Following the same rationale, the sum $Y + Z$ can be approximated by another log-normal variable, whose parameters are $\hat{\mu}_{yz}$ and $\hat{\sigma}_{yz}^2$.

Since the terms of the numerator and denominator are uncorrelated, the distribution of the ratio of log-normals is also log-normal distributed, so the resultant parameters of the distribution of the SIR are $\hat{\mu}_{xyz} = \hat{\mu}_x - \hat{\mu}_{yz}$ and $\hat{\sigma}_{xyz}^2 = \hat{\sigma}_x^2 + \hat{\sigma}_{yz}^2$. Finally, the CDF of γ_{jk}^{u}, given in (2.3), can be written as

$$\mathbf{P}[\gamma_{jk}^{\mathrm{u}} \leq \nu] = \frac{1}{2} \operatorname{erfc}\left(\frac{\log \nu - \hat{\mu}_{xyz}}{\hat{\sigma}_{xyz}^2 \sqrt{2}} \right), \tag{2.10}$$

where $\operatorname{erfc}(\cdot)$ is the complementary error function.

According to Marzetta [7], the net capacity per terminal for uplink, C_{jk}^{u}, in bits/sec/terminal, for uplink is given by

$$C_{jk}^{\mathrm{u}} = \frac{B}{\Delta} \left(\frac{T_{\mathrm{slot}} - T_{\mathrm{pilot}}}{T_{\mathrm{slot}}} \right) \left(\frac{T_{\mathrm{u}}}{T_{\mathrm{s}}} \right) \log_2(1 + \gamma_{jk}^{\mathrm{u}}), \tag{2.11}$$

where B is the total bandwidth in Hz, T_{slot} is the slot length, T_{pilot} is the time to transmit pilot sequences, T_{u} is the useful symbol duration, and T_{s} is the OFDM symbol interval, where the time is measured in seconds. Then, defining a constant $U = \frac{B(T_{\mathrm{slot}} - T_{\mathrm{pilot}})T_u}{\Delta T_{\mathrm{slot}} T_s}$, it is possible to obtain a closed-form expression for the CDF of C_{jk}^{u} as shown in

$$\mathbf{P}[C_{jk}^{\mathrm{u}} \leq c] = \frac{\left| \hat{\mu}_{xyz} - \log(2^{c/U} - 1) \right| \operatorname{erf}\left[\frac{\left| \hat{\mu}_{xyz} - \log(2^{c/U} - 1) \right|}{\sqrt{2} \hat{\sigma}_{xyz}} \right] \left| -\hat{\mu}_{xyz} + \log(2^{c/U} - 1) \right|}{2 \left| \hat{\mu}_{xyz} - \log(2^{c/U} - 1) \right|} \tag{2.12}$$

2.3.2 Infinite Number of BS Antennas

As shown in Madhusudhanan et al. [21], as the number of BS antennas is very large ($M \to \infty$), the effects of uncorrelated noise and fast fading vanish completely, and there is no interference between data transmissions inside a cell. So, the simplest linear precoders and

detectors are proved to be optimal. However, since every terminal is assigned an orthogonal time-frequency pilot sequence reused in other cells according to Δ, the only source of pilot contamination is the intercell interference. The asymptotic uplink SIR can be expressed as

$$\gamma_{jk}^{\mathrm{u}} = \frac{\zeta_{jkj}^2}{\sum\limits_{l=1,\ l\neq j} \zeta_{jkl}^2}. \tag{2.13}$$

Note that as (2.13) is composed, in the numerator by Θ, and in the denominator by S, it can be approximated by X and Y, respectively, as a ratio of two log-normal random variables whose mean and variance parameters are, respectively, given by $\hat{\mu}_{xy} = \hat{\mu}_x - \hat{\mu}_y$ and $\hat{\sigma}_{xy}^2 = \hat{\sigma}_x^2 + \hat{\sigma}_y^2$. Finally, when $M \to \infty$, the CDFs of γ_{jk}^{u} and C_{jk}^{u} can be obtained from (2.10) and (2.12), respectively, assuming $\hat{\mu}_{xyz} = \hat{\mu}_{xy}$ and $\hat{\sigma}_{xyz}^2 = \hat{\sigma}_{xy}^2$.

2.3.2.1 Non-Uniform Spatial Distribution

In this section, an approximation for (2.13) considering two types of non-uniform spatial distributions for terminal users is presented. Center-intensive user distribution and edge-intensive user distribution, which are not only analytically simple but can model practical scenarios. The center-intensive user distribution is particularly suitable for urban scenarios with populated buildings, while the edge-intensive user distribution is used to model rural mountainous propagation scenarios [22].

For center-intensive user distribution, the PDF of the distance can be expressed by

$$f_{d_{jkl}}(d) = \frac{a_c(R-d)^2 + 2d}{R^2 - r^2} b_c, \quad r \leq d \leq R \tag{2.14}$$

in which the scaling factor a_c controls how strongly the users cluster towards the BS, and $b_c = \frac{3(\mathrm{r}+\mathrm{R})}{a_c\left(\mathrm{r}^2 - 2\mathrm{rR} + \mathrm{R}^2\right) + 3(\mathrm{r}+\mathrm{R})}$. In this case, $\hat{\mu}_x$ and $\hat{\sigma}_x$ are given respectively by (2.15) and (2.16). The variable $\hat{\sigma}_x$ is a function of the parameter A.

$$\begin{aligned}\hat{\mu}_x = (r-R)\frac{-7rRa_c + r\epsilon\left(2ra_c + 9\right) + 11R^2a_c + 9R)}{3(r-R)\left[r^2a_c + r\left(3 - 2Ra_c\right) + R\left(Ra_c + 3\right)\right]} + \frac{6\mu\left(r^2a_c - 2rRa_c + R^2a_c + 3r + 3R\right)}{3(r-R)\left[r^2a_c + r\left(3 - 2Ra_c\right) + R\left(Ra_c + 3\right)\right]} \\ + \frac{6R^2\epsilon\log(R)\left(Ra_c + 3\right)}{3(r-R)\left[r^2a_c + r\left(3 - 2Ra_c\right) + R\left(Ra_c + 3\right)\right]} + \frac{6r\epsilon\log(r)\left(a_cr^2 - 3a_crR + 3a_cR^2 + 3r\right)}{3(r-R)\left[r^2a_c + r\left(3 - 2Ra_c\right) + R\left(Ra_c + 3\right)\right]}\end{aligned} \tag{2.15}$$

$$\begin{aligned}\hat{\sigma}_x = -\Big(A(\mathrm{r}-\mathrm{R})^2 - 36\mathrm{rR}^2\epsilon^2\log\left(\frac{\mathrm{r}}{\mathrm{R}}\right)\Big\{(a_c\mathrm{R}+3)\left(-3a_c\mathrm{rR} + \mathrm{r}(a_c\mathrm{r}+3) + 3a_c\mathrm{R}^2\right)\log\left(\frac{\mathrm{r}}{\mathrm{R}}\right) + \\ a_c(\mathrm{r}-\mathrm{R})[3a_c\mathrm{rR} + \mathrm{R}(9 - 5a_c\mathrm{R}) + \mathrm{r}]\Big\}\Big)^{\frac{1}{2}}\end{aligned} \tag{2.16}$$

in which

$$\begin{aligned}A = \epsilon^2\Big\{4a_c^2\mathrm{r}^4 + 2a_c\mathrm{r}^3(21 - 17a_c\mathrm{R}) + 3\mathrm{r}^2[a_c\mathrm{R}(59a_c\mathrm{R} - 26) + 27] + 2\mathrm{rR}[a_c\mathrm{R}(123 - 116a_c\mathrm{R}) + 81] + \\ \mathrm{R}^2[a_c\mathrm{R}(49a_c\mathrm{R} + 366) + 81] + 36\sigma^2\left[a_c\mathrm{r}^2 + \mathrm{r}(3 - 2a_c\mathrm{R}) + \mathrm{R}(a_c\mathrm{R} + 3)\right]^2\end{aligned} \tag{2.17}$$

On the other hand, for the edge-intensive user distribution, the PDF of the distance is given by

$$f_{d_{jkl}}(d) = \frac{a_e d^2 + 2d}{R^2 - r^2} b_e, \quad r \leq d \leq R \tag{2.18}$$

in which a_e controls how strongly the users cluster towards the cell edge and $b_e = \frac{3(\mathrm{r}+\mathrm{R})}{a_e\left(\mathrm{r}^2+\mathrm{rR}+\mathrm{R}^2\right)+3(\mathrm{r}+\mathrm{R})}$. In this case, $\hat{\mu}_x$ and $\hat{\sigma}_x$ are given by (2.19) and the variable $\hat{\sigma}_x$ is a function of the parameter B. From (2.14) and (2.18), it can be noted that, for the special case $a_c = a_e = 0$, both distributions reduce to the conventional uniform distribution.

$$\begin{aligned}
\hat{\mu}_x &= \frac{\epsilon\left[\mathrm{r}^2(2a_e\mathrm{r}+9) - 6\mathrm{r}^2(a_e\mathrm{r}+3)\log(\mathrm{r}) - 2a_e\mathrm{R}^3 + 6\mathrm{R}^2(a_e\mathrm{R}+3)\log(\mathrm{R}) - 9\mathrm{R}^2\right]}{3a_e\mathrm{r}^3 - 3\mathrm{R}^2(a_e\mathrm{R}+3) + 9\mathrm{r}^2} + 2\mu \\
\hat{\sigma}_x &= \sqrt{\frac{\epsilon^2 B}{9\left(a_e\mathrm{r}^3 - \mathrm{R}^2(a_e\mathrm{R}+3) + 3\mathrm{r}^2\right)^2} + 4\sigma^2} \\
B &= 4a_e^2\mathrm{r}^6 - 8a_e^2\mathrm{r}^3\mathrm{R}^3 + 4a_e^2\mathrm{R}^6 + 42a_e\mathrm{r}^5 - 42a_e\mathrm{r}^3\mathrm{R}^2 - 42a_e\mathrm{r}^2\mathrm{R}^3 - 36\mathrm{r}^2\mathrm{R}^2(a_e\mathrm{r}+3)(a_e\mathrm{R}+3)\log^2(\mathrm{r}) \\
&\quad -36\mathrm{r}^2\mathrm{R}^2(a_e\mathrm{r}+3)(a_e\mathrm{R}+3)\log^2(\mathrm{R}) + 36a_e\mathrm{r}^2\mathrm{R}^2(\mathrm{r}-\mathrm{R})\log(\mathrm{R}) + 36\mathrm{r}^2\mathrm{R}^2\log(\mathrm{r})[a_e(\mathrm{R}-\mathrm{r}) \\
&\quad +2(a_e\mathrm{r}+3)(a_e\mathrm{R}+3)\log(\mathrm{R})] + 42a_e\mathrm{R}^5 + 81\mathrm{r}^4 - 162\mathrm{r}^2\mathrm{R}^2 + 81\mathrm{R}^4
\end{aligned} \tag{2.19}$$

2.4 Simulation Results

For the comparison between the analytical and simulated results, it is assumed that the cellular area of interest is a tessellation filled by hexagonal cells with a radius of $r_c = 1600$ meters. The BSs are distributed deterministically in the center of each hexagon cell whose hole radius is 100 meters.

The OFDM parameters are identical to Long-Term Evolution (LTE): $T_s = 500/7\ \mu s$, $\Delta f = 15$ kHz, $T_u = 1/\Delta f$, $\frac{T_{slot}T_{pilot}}{T_{slot}} = 3/7$, and $B = 20$ MHz. The shadow-fading is modeled as a random variable, β_{jkl}, that follows a log-normal distribution with $\mu = 0$ and $\sigma = 8$ dB. The frequency reuse factor varies as $\Delta \in \{3, 7\}$, while the decay exponent is $\epsilon = 3.8$. There is no power control and the power of BSs and the user terminals are unitary, as well as γ.

Figure 2.4 shows the CDFs of SIR, in which the analytical result is represented by dashed black lines, while the solid ones refer to the simulated data. The circles indicate the curves for $\Delta \in \{3, 7\}$. As it can be seen, our approximation perfectly matches the simulated results for all range of SIR.

In order to investigate the influence of the number of antennas M, Figure 2.4 also shows the CDF for $\sigma_n = 0$, $K = 5$ and $M = 10$, in gray color, and for $M \to \infty$, in black color. Using $\Delta = 7$, for an arbitrary value of 0.6 for the CDF, it can be observed a gap of 6 dB between the SIRs. The CDF was also calculated for $M = 10^6$, and in this case, the gap, between the finite M and the asymptotic case, although small still does not vanish. So, it has been observed that

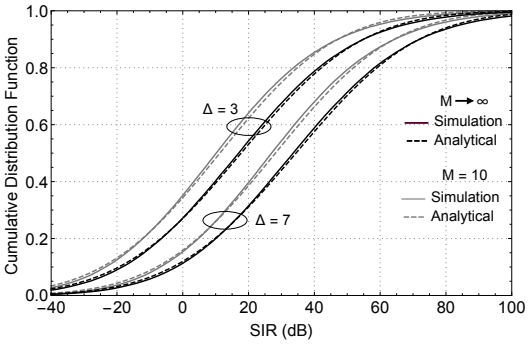


Figure 2.4 – Cumulative distribution of SIR (dB) when $M = 10$ (gray curves) and $M \to \infty$ (black curves).

the convergence between (2.3) and (2.13) is slow. In this sense, our derived expression can be used to assess how many antennas are necessary to achieve a target performance. In addition, as expected, as Δ increases, the number of interferers decreases and the SIR becomes larger, so the curve of $\Delta = 7$ is positioned to the right in this plot, while the curve of $\Delta = 3$ is in the left side.

Figure 2.5 shows the analytical (2.12) and simulated CDFs of the net uplink capacity per terminal represented by dashed and solid lines, respectively. It is observed that, for larger Δ, the capacity increases when the asymptotic SIR is low. In this region, the gains due to SIR compensate the loss by the frequency reuse, associated with the reduction in the bandwidth used by each cell. On the other side, when the SIR is high, a higher frequency reuse factor causes a net decrease in the system performance. As before, in all the cases, both simulated and analytical curves are almost indistinguishable.

Finally, considering infinite M and non-uniform user distribution, Figure 2.6 shows the CDFs of SIR, in which the analytical result is represented by dashed lines, while the solid ones refer to the simulated data. The circles indicate the curves for $\Delta = 3, 7$. The black and gray colors refer to the curves obtained from the center and edge user distributions, respectively. As it can be seen, our approximation again presents an excellent agreement.

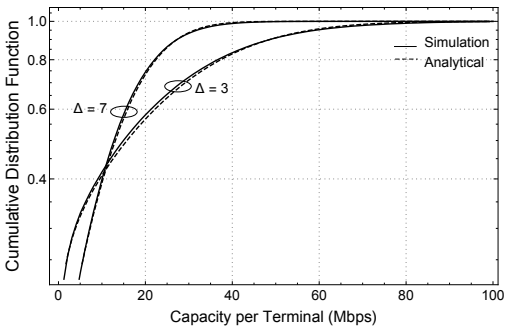


Figure 2.5 – Cumulative distribution function of the net uplink capacity per terminal (Mbps).

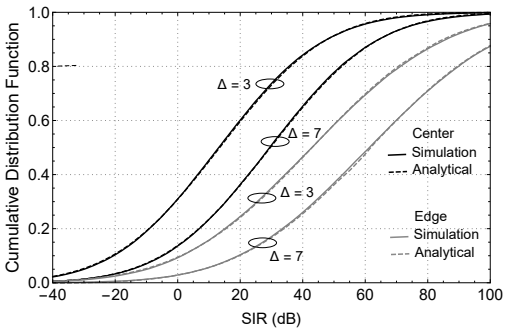


Figure 2.6 – Cumulative distribution function of SIR (dB), considering center and edge intensive user distributions.

2.5 Conclusions

In this chapter, the performance of a massive MIMO system was evaluated modeling the shadowing as a log-normal distribution. We have presented an approximated closed-form expression for the PDF and CDF of the SIR for finite and infinite number of antennas. Using the proposed approximation, the CDF for the net capacity was also derived, which allows the analysis of the system performance. All approximations are based on the moment matching method. The results have been validated by numerical simulations and have shown an excellent agreement, considering both uniform and non-uniform user distributions.

3

Preliminary Studies on the Large Intelligent Surfaces Efficiency Under Different Channels

Due to the growth in the number of applications and users requiring high-speed wireless communication, the RF spectrum is becoming overly crowded [23]. For indoor environments, the situation is even worse. The density of users and bandwidth demands are enormous, mainly thanks to the coexistence of wireless services such as cellular networks, WiFi networks, Bluetooth systems, and IoT. Therefore, to accommodate this tremendous demand for wireless, high-efficiency spectrum approaches are highly desired [24, 25].

Recently proposed, LIS is a wireless communication system that can be considered an extension of massive MIMO systems. But the idea transcends the traditional concept of antenna arrays on the base stations [26]. Its main advantage is to enable energy focus in the three-dimensional space through remote sensing with extreme precision. Because it constitutes the distributed intelligent wireless communication, sensing, and computing platform, capable of interconnecting the physical and digital worlds seamlessly and sustainably, LIS is considered the basis of future networks [27, 28].

The first studies on this subject began in 2014. From a range of experimental implementations, Hum et al. [29] showed that reconfigurable reflective arrangements and matrix lenses could dynamically control the antenna beam. A similar concept to what we now call LIS was only mentioned for the first time in 2015 at the University of California Berkeley project [15]. The general idea consisted of electromagnetically active wallpapers with built-in processing power. A chain of Field Programmable Gate Array (FPGA) automatically controls a compact integration of a large number of tiny antennas.

Nowadays, there is much research on this topic. However, challenges related to performance, user location estimation, user assignment, and power allocation need to be addressed more. Many of these works have focused on the LIS capacity for wireless communication, addressing some parameters previously developed for other systems.

For example, Björnson et al. [30] evaluate the performance of massive MIMO under adverse conditions, bringing insights on how to maximize the Energy Efficiency (EE) setting the number of BS antennas and active users and the amount of transmit power. Defined as the

number of bits transferred per Joule of energy, the EE is affected by many factors such as network architecture, transmission protocol, spectral efficiency, radiated transmit power, and circuit power consumption.

Following this same reasoning, a relevant point of the system with LIS is that it is aligned with the concepts of green communication [31]. In this topic, the critical task is to select energy-efficient communications, networking technologies, and products to minimize resource use whenever possible.

The two works of Huang et al. [3, 32] consider an LIS composed of passive antenna elements. Performing as a scatterer with reconfigurable characteristics, it does not require any dedicated energy source for either decoding, channel estimation, or transmission. The authors come across a non-convex problem based on Majorization-Maximization (MM) alternated by fractional programming to optimize the energy and spectral efficiencies. The objective function involves ratio terms. MM is an iterative optimization method that exploits a surrogate function's convexity to find their maxima or minima [33].

Although the optimization problem solution has already been given in [32], this work only considered LIS systems modeled from Rayleigh channels with Zero Forcing (ZF) precoding. Besides this, most previous research assumes that the channel fading is Rayleigh distributed. Of course, the Rayleigh-fading model is a reasonable assumption for the fading encountered in many wireless communications systems.

However, BS and LIS are part of the same infrastructure, and LISs are usually positioned to explore the Line of Sight (LoS) path concerning fixed BS in 6G networks. Hou et al. [34] use Rician fading channels for modeling the channel gain in an LIS-aided Non-Orthogonal Multiple Access (NOMA) network.

We intend to compare the result already published in the literature to what we obtained from an LoS propagation modeled from Rician channels and Matched Filter (MF). Besides, due to the high computational cost, optimization is not always possible. In this case, we derive an accurate approximation for spectral efficiency and evaluate the performance loss when BS transmits the symbols with equal powers to all users.

The remaining of this chapter is organized as follows: Section 3.1 refers to the system model, while Section 3.2 presents the problem formulation under the optimal view. Section 3.3 shows numerical simulations that confirm the superiority of Rayleigh channels in MIMO systems. An sub-optimal analysis is made in Section 3.4. Finally, we close our discussion in the Section 3.5 summarizing our conclusions.

3.1 System Model

Among a few works, there is a consensus on adopting the system model depicted as in Figure 3.1. As can be seen, the transmission occurs considering K autonomous terminals with only one antenna located in a three-dimensional environment and a two-dimensional LIS implanted in one plane. The M BS antennas transmit signals to N LIS units. Due to an obstacle, our setup consider that there is no direct path between BS and users. Also, $H_1 \in \mathbb{C}^{N \times M}$ and $H_2 \in \mathbb{C}^{K \times N}$ are the channel gain matrices between BS $\rightarrow$ LIS and LIS $\rightarrow$ k-users, respectively.

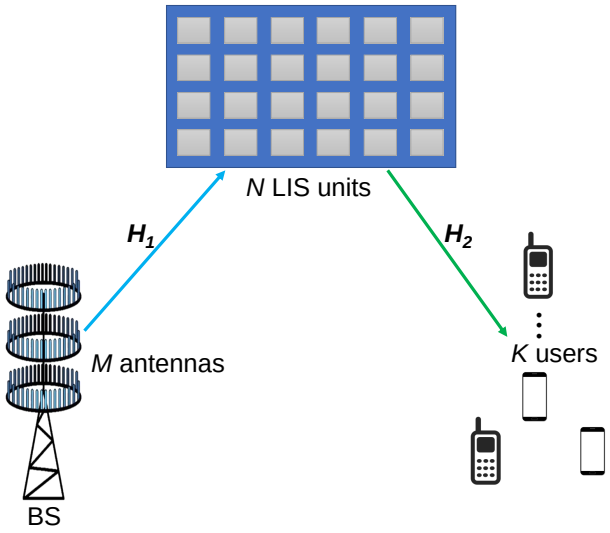


Figure 3.1 – Radiation model of a signal transmitted from BS to LIS [3].

For analytic tractability, some simplifications are made primarily concerning propagation. An ideal situation is considered with TDD protocol and perfect synchronization. Also, the BS has complete knowledge of CSI.

3.2 Problem Formulation

Considering the system model represented in Figure 3.1, the signal received at k-th mobile user can be written as

$$y_k = \beta_2 \boldsymbol{h_{2,k}} \boldsymbol{\Phi} \beta_1 \boldsymbol{H_1} x + w_k, \tag{3.1}$$

in which β_t denotes the large scale fading. In both of them, the index t refers to link 1 (BS$\rightarrow$LIS) or to link 2 (LIS$\rightarrow$ k-user). $h_{2,k} \in \mathbb{C}^{1 \times N}$ is the channel gain vector between LIS and k-user and represent the small scale fading. Also, $\boldsymbol{\Phi}$ is a diagonal matrix of dimension $N \times N$ whose elements $\Phi = e^{j\theta_n}$ are the effective phase shifts applied by all LIS reflecting elements, in which $\theta_n \in [0, 2\pi]$. $w_k \sim \mathbb{CN}(0, \sigma^2)$ is the thermal noise modeled as a realization of a zero-mean

complex circularly symmetric Gaussian variable with variance σ_w^2. Finally, $x = \sum_{k=1}^{K} \sqrt{p_k} g_k s_k$. In this expression p_k, g_k and s_k denote the transmit power, the precoding vector, and the unit power complex-valued information symbol, respectively.

Then, the Signal-to-Interference-plus-Noise Ratio (SINR) is given by

$$\gamma_k = \frac{\beta_k p_k |\boldsymbol{h_{2,k}} \boldsymbol{\Phi} \boldsymbol{H_1} g_k|^2}{\sum_{i=1, i\neq k}^{K} \beta_i p_i |\boldsymbol{h_{2,k}} \boldsymbol{\Phi} \boldsymbol{H_1} g_i|^2 + \sigma_w^2}, \tag{3.2}$$

in which β_k is the k-user path-loss obtained by the product of β_1 and β_2.

The EE is defined as the ratio between the system achievable sum rate in bps and the total power consumption in Joule. The optimization problem appears from (3.2), since it is necessary to define the transmit powers for all users and the values for the LIS elements that jointly maximize the bit-per-Joule energy efficiency performance. Therefore, the problem consists in to solve the following non-convex optimization [35]:

$$\begin{aligned} \underset{bm\Phi, P}{\text{maximize}} \quad & \frac{\sum_{k=1}^{K} \log_2(1+\gamma_k)}{\xi \sum_{k=1}^{K} p_k + P_{BS} + K P_{UE} + N P_n} \\ \text{subject to} \quad & \log_2(1+\gamma_k) > R_{min,k}, \ \forall k = 1, \ldots, K, \\ & \beta^2 \, \text{Tr}[\boldsymbol{G P G}^H] \leq P_{max}, \\ & |\Phi| = 1 \ \forall n = 1, \ldots, N, \end{aligned} \tag{3.3}$$

where ^{H}, and Tr indicate Hermitian (conjugate transpose), and trace of a matrix, respectively. In its turn, ξ is the efficiency of the transmit power amplifier, P_{UE} is the hardware static power dissipated by the k-th user while P_{BS} and P_n denote the total hardware static power consumption at BS and LIS, respectively. $R_{min,k}$ denotes the individual QoS constraint of the k-th user, P_{max} is the maximal system power and $\boldsymbol{P} = \text{diag}[p_1, \ldots, p_k] \in \mathbb{C}^{K \times K}$.

If we consider only the numerator of the objective function on (3.3), it can also optimize the Spectral Efficiency (SE). This parameter measures the information rate transmitted over a given bandwidth in a specific communication system. That is, how efficiently a limited frequency spectrum is used. Its optimization problem consists of (3.3) considering $\xi = 0$.

Our goal here is to analyze the LIS-assisted system under different scenarios. Therefore, in relation to g_k:

- It can be Zero Forcing (ZF) precoding whose matrix is given by $\boldsymbol{G} = (\beta_2 \boldsymbol{H}_2 \boldsymbol{\Phi} \beta_1 \boldsymbol{H}_1)^+$, in which $^+$ means pseudo-inverse, or
- It can be Matched Filter (MF) precoding whose matrix is given by $\boldsymbol{G} = (\beta_2 \boldsymbol{H}_2 \boldsymbol{\Phi} \beta_1 \boldsymbol{H}_1)^H$.

It is worth mentioning that in both cases, $\boldsymbol{H}_2 = [\boldsymbol{h_{2,1}^T}, \boldsymbol{h_{2,2}^T}, \ldots, \boldsymbol{h_{2,K}^T}]^T \in \mathbb{C}^{K \times N}$. Regarding the channel, both $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ can be modeled according to:

- Rayleigh fading: characterized by multipaths and modeled as $[\boldsymbol{H_1}]_{i,j} \sim \mathbb{CN}(0, \sigma^2)$ and $[\boldsymbol{h}_{2,k}]_i \sim \mathbb{CN}(0, \sigma^2)]$, or
- Rician fading: when typically a dominant line of sight signal is much stronger than the others. In this case, $[\boldsymbol{H_1}]_{i,j} \sim \mathbb{CN}(K_F, \sigma^2)]$ and $[\boldsymbol{h_{2,k}}]_i \sim \mathbb{CN}(K_F, \sigma^2)]$, where the shape parameter, K_F, is deterministic and is defined as the ratio of the power contributions by LoS path to the remaining multipaths.

It is worth mentioning that high processing resource is not always possible. There is a hardware limitation. In these situations, the best thing to do is to assume equal power for all users. Therefore, a sub-optimal analysis is also presented in Section 3.4.

3.3 Numerical Results

In this section, numerical results are presented to validate the simulations obtained from 10^3 realizations, considering the practical values shown in Table I. The efficient MM approach detailed in [32] is adopted. Beyond the variables defined in the table, it is considered that both, $\boldsymbol{h_{2,k}}$ and $\boldsymbol{H_1}$, follow Rayleigh or Rician distribution and $K_F \in \{1, 2, 4\}$. The multiple single-antenna mobile users are assumed randomly and uniformly placed in the 100m×100m half right-hand side rectangular, according to Figure 3.2.

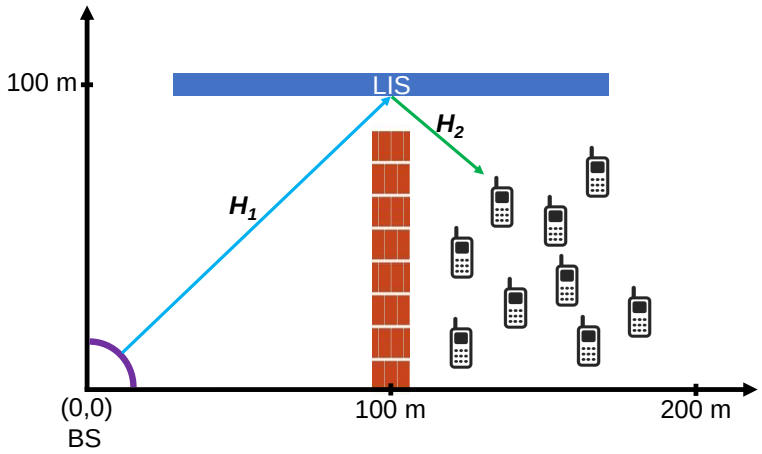


Figure 3.2 – Scenario adopted with K uniformly distributed users [3].

Figure 3.3 compares the optimization technique performance under different propagation conditions from the EE point of view. Two curves refer to the MM-based method considering the Rayleigh model, while the others are related to LoS with $K_F = 1$. Note that the channel's energy efficiency under the Rayleigh fading has a higher energy efficiency than the Rician case. The reason for this is the well studied rank degradation in MIMO environments and its reduction on the spectral efficiency [36].

Table 3.1 – Parameters adopted for the Monte Carlo simulations.

Description	**Value**
Noise variance, σ^2	1
Circuit dissipated power coefficients at BS, ξ	1.2
Large scale fading, β_t	$10^{-3.53}/d_t^{3.76}$
Circuit dissipated power, P_{BS}	40 dBm
Dissipated power at each user, P_{UE}	10 dBm
Dissipated power at n-th LIS element, P_n	10 dBm
Individual QoS requirements, $R_{min,k}$	0
Maximum transmit power at BS, P_{max}	50 dBm
Transmission bandwidth	180 kHz

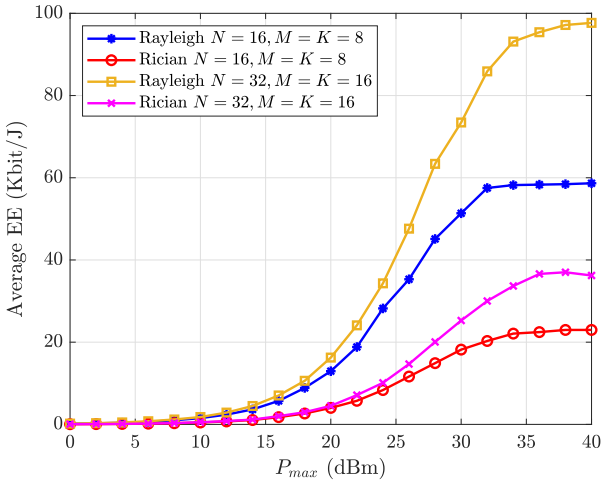


Figure 3.3 – Average EE of Rician and Rayleigh channels under different setups.

Figure 3.4 presents this same parameter but considering the fading modeled as Rician distribution with different K_F deterministic factors. It is possible to notice that the higher K_F, the higher the power of the dominant component and the worse is the system performance when the channel is modeled as Rice distribution, consequently.

In its turn, Figure 3.5 compares the MF and ZF performances considering the SE of Rayleigh and Rician channels. Just as in conventional MIMO systems, ZF stands out for Rayleigh channels. However, when we consider the LoS scenario, MF precoding presents better performance.

Finally, Figure 3.6 compares the EE obtained from MF and ZF techniques for two different schemes with the Rayleigh channel. As expected, MF behaves better for low SNRs while the performance of ZF stands out for high SNRs.

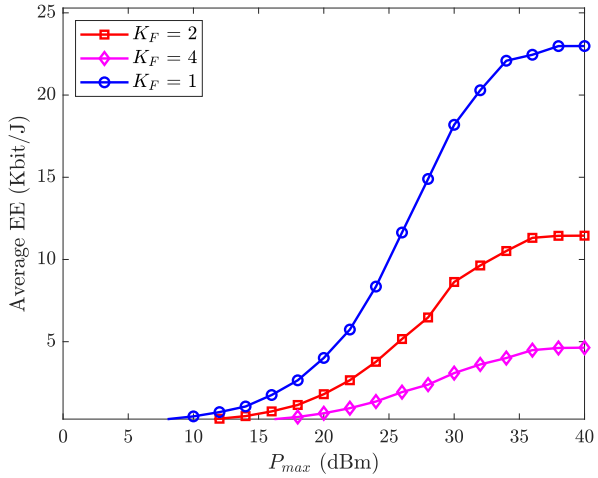


Figure 3.4 – Average EE of Rician channels with different K_F factors and $K = 8, M = 8, N = 16$.

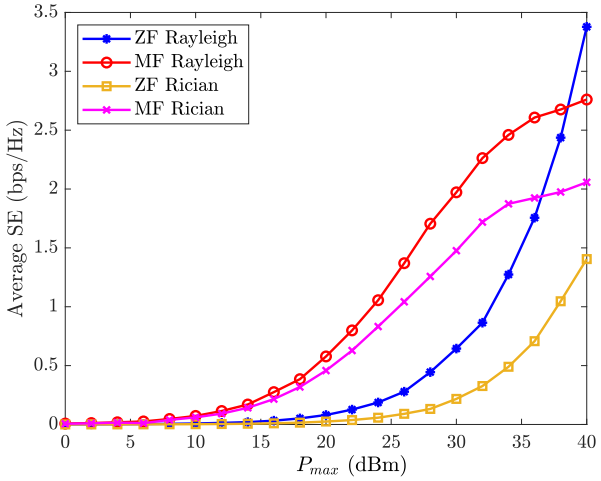


Figure 3.5 – Comparison between ZF and MF applied to Rayleigh and Rician channels with $N = 16$, $K = 8$ and $M = 8$.

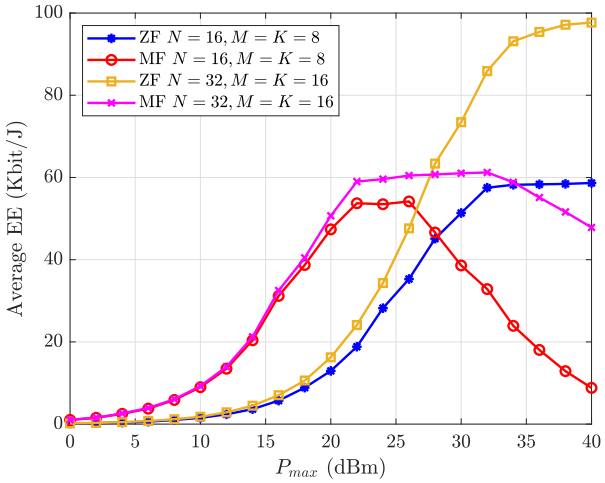


Figure 3.6 – EE considering ZF and MF techniques for Rayleigh channels.

3.4 Equal Power ZF

Another interesting aspect that deserves to be verified is to measure the loss in the system performance when the total power is not optimized. That is because it is not always possible to have the channel state information in real-time of a given effective system. In these cases, the best we can do is assume the uniform transmit power for all K users. So, the condition in (3.3) becomes an equality and we can write $\mathrm{Tr}(GPG^H) = P_{max}$. Then

$$p_k = \frac{P_{max}}{K\,\mathrm{Tr}(\alpha)}, \tag{3.4}$$

in which $\alpha = (\beta_2 H_2 \boldsymbol{\phi} \beta_1 H_1)(\beta_2 H_2 \boldsymbol{\phi} \beta_1 H_1)^H$. For simplicity, let $\beta = \beta_1\beta_2$ from now on.

Considering ZF precoding, let the sum-capacity be given by the sum of the spectral efficiency of all users as

$$\begin{aligned} C &= \sum^{K} \log_2(1 + p_k \sigma_w^{-2}) \\ &= K \log_2\left(1 + \frac{P_{max}}{KTr[\alpha]}\right). \end{aligned} \tag{3.5}$$

Then, it is possible to obtain an analytical expression from the following steps.

Since α is a random variable that depends on path loss represented by β, the probability function is $p(\alpha, \beta) = p(\alpha|\beta)p(\beta)$ and so the expected value for the sum-capacity can be expressed by

$$\bar{C} = K \int_0^\infty \int_0^\infty \log_2\left(1 + \frac{P_{max}}{KTr[\alpha]}\right) p(\alpha|\beta)p(\beta)d\alpha d\beta. \tag{3.6}$$

A possible simplification is to replace β for its expected value. Then

$$\mathbb{E}[\mathrm{Tr}(\alpha)] = \mathbb{E}[\beta^2]\mathbb{E}[\mathrm{Tr}(\alpha_1)], \tag{3.7}$$

and

$$\mathrm{var}[\mathrm{Tr}(\alpha)] = \mathbb{E}[\beta^2]\sqrt{\mathrm{var}[\mathrm{Tr}(\alpha_1)]} \tag{3.8}$$

in which $\alpha_1 = (H_2 \boldsymbol{\phi} H_1)(H_2 \boldsymbol{\phi} H_1)^H$ and var(.) means variance operator. In this case, (3.6) can be rewritten by

$$\bar{C} = K \int_0^\infty \log_2\left(1 + \frac{P_{max}}{\mathbb{E}[\beta^2]KTr[\alpha_1]}\right) p(\alpha_1)d\alpha_1. \tag{3.9}$$

From the central limit theorem, we can assume that $p(\alpha_1)$ follows a normal distribution. In this case, the mean value can be obtained as

$$\begin{aligned} \mathbb{E}[\mathrm{Tr}(\alpha_1)] &= \sum_{i=1}^{K}\sum_{i=1}^{M} \mathbb{E}[|g_{i,j}|^2] \\ &= KMN\sigma^2\sigma_\phi^2\sigma^2 \\ &= KMN, \end{aligned} \tag{3.10}$$

in which $g_{i,j}$ is the element of $\mathbf{G}$ located in the i-th row and in the j-th column. σ_{ϕ}^2 is the variance of $\boldsymbol{\phi}$. On the other hand, the second moment $\mathbb{E}[\mathrm{Tr}^2(\alpha_1\alpha_1^H)]$ can be simplified by the following expression using Mathematica software [37]:

$$\begin{aligned}\mathbb{E}[\mathrm{Tr}^2(\alpha_1)] &= \sum_{b=1}^{N}\sum_{r=1}^{M}\sum_{p=1}^{K}\sum_{t=1}^{K}\sum_{b_1=1}^{N}\sum_{r_1=1}^{M}\sum_{p_1=1}^{K}\sum_{t_1=1}^{K}\boldsymbol{\Phi} \\ &= KMN(K+M+N+KMN) \end{aligned} \tag{3.11}$$

in which Φ is expressed below as

$$\Phi = \begin{cases} 1, \text{ if } (b=b_1)\cap(r=r_1)\cap(p_1=t)\cap(p=t_1)\cap(p\neq t)\cup(p_1\neq t_1)\cup(p=t\cap(p_1=t_1)\cap(b\neq b_1)\cap(r\neq r_1)\cup(p\neq p_1) \\ 2, \text{ if } (b\neq b_1)\cup(r\neq r_1)\cap(b=b_1)\cup(r=r_1)\cap(p=p_1)\cap(p=t)\cap(p_1=t_1) \\ 4, \text{ if } (b=b_1)\cap(p=p_1)\cap(r=r_1)\cap(p=t)\cap(p_1=t_1) \end{cases} \tag{3.12}$$

Since $\mathrm{var}[\mathrm{Tr}(\alpha_1)] = \mathbb{E}[\mathrm{Tr}^2(\alpha_1)] - \mathbb{E}^2[\mathrm{Tr}(\alpha_1)]$ [18], then

$$\mathrm{var}[\mathrm{Tr}(\alpha_1)] = KMN(K+M+N) \tag{3.13}$$

Comparatively, Figure 3.7 shows three versions of sum-capacity: one optimum simulated obtained by optimizing the transmit power and effective phase shifts of each user, one simulated obtained through (3.6), and finally, one approximation. It is worth emphasizing that these last two are sub-optimal since they consider the same transmit power for all users. The logarithmic scale is adopted to make clear the tendency of the curves. It is evident the MM optimization technique robustness since it provides a gain of 6 dB when used. However, considering situations in which CSI is not available, the approximation derived here is very accurate concerning the simulated sub-optimal results.

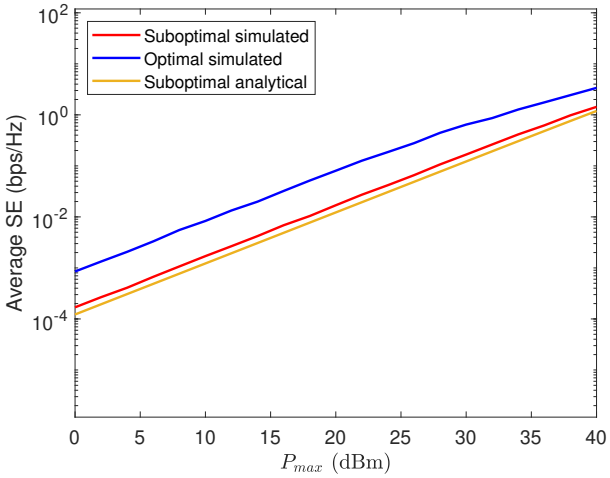


Figure 3.7 – Comparison between optimal, and suboptimal sum-capacity for $K = 8$, $M = 8$, $N = 16$.

3.5 Conclusions

In this chapter, we have reproduced some LIS-assisted systems modeled from Rayleigh fading and ZF precoding. Then, they are compared to systems that consider Rician channels and MF precoding. As already mentioned in the literature [38,39], we have confirmed that Rayleigh channels perform better than Rician ones even in not conventional MIMO environments, like LIS.

This result can be explained from the spread, the higher the scattering, the better the system performance. What increases the spread, in this case, is the number of LIS reflecting units and the fact that the channel is Rayleigh composed of various multipath.

Besides, ZF is better when the system power is high. However, in adverse environments, when it is low, it is clear the MF superiority. A suboptimal analysis has also been made assuming the same power for all users. An average gain of 6 dB in sum-capacity can be observed when the optimization is applied.

4

Realistic Large Intelligent Surface-Assisted Systems

5G wireless technology is expected to be the most reliable and lowest latency technology ever. It also looks to deliver massive connectivity to many devices that require low bandwidth and low power. However, there is no single physical layer technology that can satisfy all requirements of 5G. To overcome this issue, one of the solutions of the 6G wireless technology is to exploit the broadcast nature of wireless propagation intelligently. The advent of LIS enables network operators to intelligently control the radio propagation environment in order to boost the signal quality [40].

At first glance, it may seem a futuristic technology the idea of a smart wireless communication capable of self-reconfiguring and sensing, processing, and learning from the environment characteristics. Actually, this is already a reality and is being tested and marketed [41, 42]. This system consists of many nearly passive antenna elements, whose parameters can be tuned according to desired objectives. It is also able to be reconfigured based on the information received from the environment [27].

As described in the works [32, 43], there are some similarities of operation between LIS and amplify-and-forward (AF) relaying. However, it is worth mentioning that, although several studies already model the performance of AF systems [44, 45], the same techniques cannot be used for LIS. In addition to the fact that there is no power amplifier in an LIS, its reflection coefficients are complex entities with varying amplitude and phase.

Thus, it is essential to model the performance focusing on LIS-assisted systems. Although it is a recent area of study, well-known works are highlighting the advantages of such technology.

Under hostile conditions like estimation errors and interference channels, Jung et al. [46] evaluate the fundamental performance benefits of practical LIS environments, comparing it to that obtained in massive multiple-input multiple-output (MIMO) systems. They conclude that disturbs, like hardware impairments, noise, and interference from estimation errors, become negligible as the number of LIS elements increases. A LIS-assisted system can also achieve a performance comparable to conventional massive MIMO with improved reliability and a significantly reduced area for antenna deployment.

Focusing on the link budget path-loss versus error performance, Basar et al. [47] models

a perfect LIS and proves that the received power does decay with the second power of the distance, just like the LoS ray. Considering phase noise, Badiu et al. [48] analyze such a system's performance and shows its robustness even under imperfect phase estimation or phase quantization with errors. Although interesting, all these works mentioned above consider the ideal scenario in which the reflection coefficient amplitude is fixed and equal to one.

From Yuen et al. [49], we know now that the phase shift causes reflective currents capable of altering the previously reflection amplitude. However, the consequences of these changes for the total system performance remain unclear.

This chapter aims to analyze the performance of single-input single-output (SISO) LIS-assisted systems taking into account a non-unitary reflection coefficient with and without phase error. This analysis provides insight into the performance loss inherent in practical systems. More specifically, we characterize the equivalent channel between the source and the destination. As the number of LIS elements increases, we can apply the central limit theorem (CLT). As a consequence, the in-phase and quadrature components of the complex channel response can be approximated by a Gaussian random variable. With this assumption, we assume a binary phase-shift keying (BPSK) modulation and derive the exact and a very tight upper bound for the symbol error probability.

The remaining of this chapter is organized as follows: Section 4.1 refers to the system model, while Section 4.2 presents the mathematical formulation and analytical expressions obtained for the symbol error probability in different cases. Section 4.3 shows numerical comparisons that confirm the accuracy of our approximation. Finally, we close the discussion in Section 5.5 summarizing our conclusions.

4.1 System Model

The considered SISO wireless system comprises a single-antenna source (S), a single-antenna destination (D), and an LIS composed for n reflective elements, as shown in Figure 4.1. These structures, identified with index $i = 1 \dots n$, are responsible for assisting the communication between S and D since the direct link between them is so hostile, it can be overlooked.

Assuming slow and flat fading, let H_{i1} and H_{i2} be the complex fading coefficients of the S-to-R$_i$ and R$_i$-to-D channels, respectively. Also, we consider that the reflectors are sufficiently spaced apart, such that all the coefficients are mutually independent.

Then, the signal received by the destination over the equivalent baseband channel is

$$y = n\sqrt{E_s}|H|x + w, \tag{4.1}$$

in which E_s is the average transmitted energy per symbol, x is the binary transmitted signal ($x \in \{-1, 1\}$), and w is the additive white Gaussian noise with variance σ_w^2.

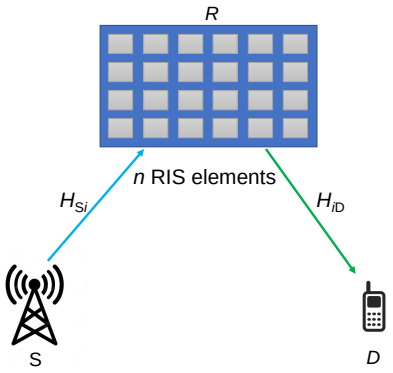


Figure 4.1 – System model with LIS between S and D.

The channel response can be written as

$$H = \frac{1}{n}\sum_{i=1}^{n} |H_{i1}|\,|H_{i2}|\,\beta(\theta_i)e^{\theta_i-\alpha_i-\delta_i} \quad \in \mathbb{C}, \tag{4.2}$$

in which $\alpha_i = \measuredangle H_{i1}$ and $\delta_i = \measuredangle H_{i2}$ are the phase of both links. Also, let $\beta(\theta_i)e^{\theta_i} \in \mathbb{C}^{n\times 1}$ denote the reflection coefficient vector of the LIS ith element. $\beta(\theta_i) \in [0,1]$ and $\theta_i \in [-\pi,\pi]$ are the amplitude and the phase shift on the combined incident signal, respectively. As defined in [49], when the phase shift is around π or $-\pi$, the reflective currents are out-of-phase with the element currents. Thus the electric field and the current flow in the element are diminished, resulting in the highest reflection amplitude. In contrast, when the phase shift is around zero, the reflective currents are in-phase with the element currents. The electric field and the current flow in the element are both enhanced. As a result, the dielectric loss, metallic loss, and ohmic loss increase dramatically, leading to substantial energy dissipation and the lowest reflection amplitude. So,

$$\beta(\theta_i) = \beta_{\min} + (1-\beta_{\min})\left\{\frac{1}{2}[\sin(\theta_i-\phi)+1]\right\}^k, \tag{4.3}$$

in which $\beta_{\min} \geq 0$, $\phi \geq 0$, and $k \geq 0$ are the constants related to the specific circuit implementation. $\beta_{\min}$ is the minimum amplitude, ϕ is the horizontal distance between $-\frac{\pi}{2}$ and $\beta_{\min}$, and k controls the steepness of the function curve. Note that for the ideal phase shift model $|\beta(\theta_i)| = 1,\ \forall n$, regardless of the phase shift.

4.2 Symbol Error Probability Approximation

This section discusses the different forms the exponent in (4.2) can take and the respective mathematical implications. Therefore, there are four possible cases, and it is worth mentioning that in all of them, θ_i is assumed uniformly distributed without loss of generality.

4.2.1 Case I ($\beta(\theta_i) = 1$ and $\theta_i = \alpha_i + \delta_i$)

As stated in [47], from (4.1) we have the following instantaneous SNR in D

$$\gamma = \frac{|\sum_{i=1}^{n} |H_{i1}|\,|H_{i2}|\,\beta(\theta_i) e^{\theta_i - \alpha_i - \delta_i}|^2 E_s}{\sigma_W^2}, \tag{4.4}$$

whose maximal value happens when $\theta_i = \alpha_i + \delta_i$ and so there is no phase error. Assuming that the reflection coefficient amplitude is unitary, this is the optimal case. Its symbol error probability for BPSK modulation is given by [47, Eq. 17]

$$P_e = \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \sqrt{\frac{1}{1 + \frac{(16-\pi^2)E_s n}{8\sigma_W^2 \sin^2(\eta)}}} \exp\left(- \frac{\pi^2 E_s n^2}{[16\sigma_W^2 \sin^2(\eta)]\left[\frac{(16-\pi^2)E_s n}{8\sigma_W^2 \sin^2(\eta)} + 1\right]} \right) d\eta, \tag{4.5}$$

and the corresponding upper bound is given by [47, Eq. 18]

$$P_e \leq \frac{1}{2} \sqrt{\frac{1}{1 + \frac{(16-\pi^2)E_s n}{8\sigma_W^2)}}} \exp\left(- \frac{\pi^2 E_s n^2}{[16\sigma_W^2)]\left[\frac{(16-\pi^2)E_b n}{8\sigma_W^2} + 1\right]} \right). \tag{4.6}$$

4.2.2 Case II ($\beta(\theta_i) = 1$ and $\theta_i \neq \alpha_i + \delta_i$)

On the other hand, when the amplitude of the reflection coefficient remains unitary, but there are phase errors, the result in [48] can be used. According to it, the phase error represented by uniform distribution means high uncertainty. Then, $\phi_p = 0$ for $p \geq 1$ and we have that H follows a normal distribution. Its real and imaginary parts, $C = \mathbb{R}(H)$ and $S = \mathbb{I}(H)$ respectively, are independent, with mean $\mu_C = \mu_S = 0$ and variance $\sigma_C^2 = \sigma_S^2 = \frac{1}{2n}$. Then, the symbol error probability is approximated by

$$P_e = \int_0^{2\pi} \int_0^{\infty} Q\left(\sqrt{\frac{n^2 E_s \gamma}{2\sigma_W^2}} \right) \times \frac{\exp\left[-\frac{\gamma \cos^2(\zeta)}{2\sigma_C^2} - \frac{\gamma \sin^2(\zeta)}{2\sigma_S^2} \right]}{4\pi \sigma_C \sigma_S} d\gamma d\zeta, \tag{4.7}$$

while its upper bound is given by

$$P_e \leq \frac{\sqrt{n}\sigma_C \sigma_W^2}{2\sqrt{n}\sigma_C \left(E_s n^2 \sigma_S^2 + \sigma_W^2\right)}. \tag{4.8}$$

4.2.3 Case III ($\beta(\theta_i) \neq 1$ and $\theta_i = \alpha_i + \delta_i$)

Here, we consider that the reflection coefficient amplitude is different from one, and there is no error concerning estimation and quantization process. Just like in [47], the phase θ_i is set to cancel the overall phase shift, which maximizes the SNR at the receiver [50]. However, even so, θ_i causes a variation in the reflection coefficient amplitude. Then, the channel response becomes

$$H = \frac{1}{n} \sum_{i=1}^{n} |H_{i1}|\,|H_{i2}|\,\beta(\theta_i). \tag{4.9}$$

As n tends to infinity, it is known that H can be treated as a Gaussian random variable distributed. Therefore, our approach proposes to find the mean and variance parameters of this distribution, and then the symbol error rate as follows.

4.2.3.1 Mean of H

We know that the random variables are independent of each other and H_{i1} and H_{i2} are Rayleigh variables defined by σ_R parameter. Let the term μ_β represents the expectation of $\beta(\theta_i)$, then

$$\mu_\beta = \beta_{\min} - \frac{(\beta_{\min} - 1)\Gamma\left(k + \frac{1}{2}\right)}{\sqrt{\pi}\Gamma(k+1)}, \tag{4.10}$$

in which $\Gamma[.]$ represents the Gamma function. Then, substituting this expression, the mean of H, μ_H, can be obtained as

$$\mu_H = \frac{1}{2}\pi\sigma_R^2\left[\beta_{\min} - \frac{(\beta_{\min} - 1)\Gamma\left(k + \frac{1}{2}\right)}{\sqrt{\pi}\Gamma(k+1)}\right]. \tag{4.11}$$

4.2.3.2 Variance of H

Regarding the variance of $\beta(\theta_i)$, it can be computed as

$$\sigma_\beta^2 = \frac{(\beta_{\min} - 1)^2\left[\frac{\sqrt{\pi}\Gamma\left(2k+\frac{1}{2}\right)}{\Gamma(2k+1)} - \frac{\Gamma\left(k+\frac{1}{2}\right)^2}{\Gamma(k+1)^2}\right]}{\pi}, \tag{4.12}$$

and so we have that the variance of H is given by

$$\sigma_H^2 = \frac{\sigma_R^4\left(\frac{(\beta_{\min}-1)\left(-\frac{\pi^{3/2}(\beta_{\min}-1)\Gamma\left(k+\frac{1}{2}\right)^2}{\Gamma(k+1)^2} + \frac{2\left(\pi^2-16\right)\beta_{\min}\Gamma\left(k+\frac{1}{2}\right)}{\Gamma(k+1)} + \frac{16(\beta_{\min}-1)\Gamma\left(2k+\frac{1}{2}\right)}{\Gamma(2k+1)}\right)}{\sqrt{\pi}} - (\pi^2-16)\,\beta_{\min}^2\right)}{4n} \tag{4.13}$$

4.2.3.3 Symbol Error Probability

Defined the mean and variance of H, we can then say that the SNR is a Gaussian random variable with the following probability density function [19]

$$f(\gamma) = \frac{e^{-\frac{(\mu_H+\sqrt{\gamma})^2}{2\sigma_H^2}}\left(e^{\frac{2\mu_H\sqrt{\gamma}}{\sigma_H^2}} + 1\right)}{2\sqrt{2\pi}\sigma_H\sqrt{\gamma}}. \tag{4.14}$$

Considering the BPSK modulation, it is simple to show that the symbol error probability is given by [51]

$$P_e = \int_0^{\infty} Q\left(\sqrt{\frac{n^2 E_s \gamma}{2\sigma_W^2}}\right) \frac{\left[e^{-\frac{(\mu_H+\sqrt{\gamma})^2}{2\sigma_H^2}}\left(e^{\frac{2\mu_H\sqrt{\gamma}}{\sigma_H^2}}+1\right)\right]}{2\sqrt{2\pi}\sigma_H\sqrt{\gamma}} d\gamma, \tag{4.15}$$

in which $Q(\cdot)$ is the Gaussian error function and can be approximated from the Chernoff bound as $Q(x) \leq \frac{1}{2}e^{-\frac{x^2}{2}},\ x > 0$. Then, the following upper bound can be found from (4.15)

$$P_e \leq \frac{e^{-\frac{E_s\mu_H^2 n^2}{2\left(E_s n^2\sigma_H^2+\sigma_W^2\right)}}}{2\sigma_H\sqrt{\frac{E_s n^2}{\sigma_W^2}+\frac{1}{\sigma_H^2}}}, \tag{4.16}$$

whose limit is given by $\lim_{n\to\infty} P_e = 0$ when n tends to infinity.

4.2.4 Case IV ($\beta(\theta_i) \neq 1$ and $\theta_i \neq \alpha_i + \delta_i$)

On the other hand, Case IV considers that there is an error related to the estimation or quantization of phase. Assuming that the user on the destination is fixed, and the phases α_i and δ_i do not vary over time, set $\phi = 0$. Then H can be written as

$$\begin{aligned} H &= \frac{1}{n}\sum_{i=1}^{n} |H_{i1}|\,|H_{i2}|\,\beta(\theta_i)e^{\theta_i-\alpha_i-\delta_i} \\ &= \frac{1}{n}\sum_{i=1}^{n} |H_{i1}|\,|H_{i2}|\,\beta(\theta_i)[\cos(\theta_i-\alpha_i-\delta_i)+j\sin(\theta_i-\alpha_i-\delta_i)]. \end{aligned} \tag{4.17}$$

To deal with the approximation, its absolute square value is given by

$$|H|^2 = \frac{1}{n^2}\Bigg\{\Bigg[\sum_{i=1}^{n} |H_{i1}|\,|H_{i2}|\,\beta(\theta_i)\cos(\theta_i-\alpha_i-\delta_i)\Bigg]^2 + \Bigg[\sum_{i=1}^{n} |H_{i1}|\,|H_{i2}|\,\beta(\theta_i)\sin(\theta_i-\alpha_i-\delta_i)\Bigg]^2\Bigg\}. \tag{4.18}$$

Following the same reasoning mentioned above to calculate the mean and variance, the definition below is considered from now on. Let

$$\begin{aligned} C &= \frac{1}{n}\sum_{i=0}^{n} |H_{i1}|\,|H_{i2}|\,\beta(\theta_i)\cos(\theta_i-\alpha_i-\delta_i), \\ S &= \frac{1}{n}\sum_{i=0}^{n} |H_{i1}|\,|H_{i2}|\,\beta(\theta_i)\sin(\theta_i-\alpha_i-\delta_i). \end{aligned} \tag{4.19}$$

4.2.4.1 Mean of C and S

Be the mean of $\beta(\theta_i)\cos(\theta_i-\alpha_i-\delta_i)$ defined as μ_{β_C}, we know that the media of C is then

$$\begin{aligned}\mu_C &= \frac{1}{n}\sum_{i=1}^{n} E[H_{i1}H_{i2}]\mu_{\beta_C} \\ &= \frac{1}{n}\sum_{i=1}^{n} -2^{-k-3}\sigma_R^2\left(\beta_{\min}-1\right)\sin\left(\alpha_i+\delta_i\right) \\ &\quad\left[-\frac{4_3F_2\left(\frac{1}{2},1,\frac{3}{2};\frac{k}{2}+\frac{3}{2},\frac{k}{2}+2;1\right)}{k^2+3k+2}+\right. \\ &\quad\left. 4_3F_2\left(1,\frac{1}{2}-\frac{k}{2},-\frac{k}{2};\frac{1}{2},\frac{3}{2};1\right)+\pi k_2F_1\left(\frac{1-k}{2},1-\frac{k}{2};2;1\right)\right], \quad (4.20)\end{aligned}$$

in which ${}_3F_2(.)$ is the generalized hypergeometric function and ${}_2F_1(.)$ is the hypergeometric function. Similarly, the mean of S becomes

$$\begin{aligned}\mu_S &= \frac{1}{n}\sum_{i=1}^{n} E[H_{i1}H_{i2}]\mu_{\beta_S} \\ &= \frac{1}{n}\sum_{i=1}^{n}\left(\beta_{\min}-1\right)\left(-2^{-k-3}\right)\sigma_R^2\cos\left(\alpha_i+\delta_i\right) \\ &\quad\left[-\frac{4_3F_2\left(\frac{1}{2},1,\frac{3}{2};\frac{k}{2}+\frac{3}{2},\frac{k}{2}+2;1\right)}{k^2+3k+2}+\right. \\ &\quad 4_3F_2\left(1,\frac{1}{2}-\frac{k}{2},-\frac{k}{2};\frac{1}{2},\frac{3}{2};1\right) \\ &\quad\left.+\pi k_2F_1\left(\frac{1-k}{2},1-\frac{k}{2};2;1\right)\right]. \quad (4.21)\end{aligned}$$

4.2.4.2 Variance of C and S

We know that the variance of β_C and β_S can be found as $\sigma^2_{\beta_C} = E[\beta_C(\theta)^2] - \mu^2_{\beta_C}$ and $\sigma^2_{\beta_S} = E\left[\beta_S(\theta)^2\right] - \mu^2_{\beta_S}$, respectively. From a straightforward way, we can obtain σ_C^2 and σ_S^2 as expressed in Appendix A.

4.2.4.3 Symbol Error Probability

Since $|H|^2$ can be written as $|H|^2 = C^2+S^2$, the distribution of $R=|H|$ can be computed from the distributions of C and S as [19]

$$f_{R,Z}(r,\zeta) = r f_{C,S}(c,s). \quad (4.22)$$

Since C and S are Gaussian independent random variables, the following can be written

$$f_{R,Z}(r,\zeta) = r\frac{1}{2\pi\sigma_C\sigma_S}e^{-\frac{(C-\mu_C)^2}{2\sigma_C^2}-\frac{(S-\mu_S)^2}{2\sigma_S^2}}. \quad (4.23)$$

Also, $C = R\cos(\zeta)$ and $S = R\sin(\zeta)$, then we have that

$$f_{R,Z}(r,\zeta) = \frac{r}{2\pi\sigma_C\sigma_S} e^{-\frac{(r\cos(\zeta)-\mu_C)^2}{2\sigma_C^2} - \frac{(r\sin(\zeta)-\mu_S)^2}{2\sigma_S^2}}. \tag{4.24}$$

Defining the variable $\gamma = |H|^2 = R^2$, the joint distribution of γ and θ can be obtained through a variable transformation [19] and the joint probability density function of γ and θ can be computed as

$$f(\gamma,\zeta) = \frac{\exp\left\{-\frac{\left[\mu_C-\sqrt{\gamma}\cos(\zeta)\right]^2}{2\sigma_C^2} - \frac{[\mu_S-\sqrt{\gamma}\sin(\zeta)]^2}{2\sigma_S^2}\right\}}{4\pi\sigma_C\sigma_S}. \tag{4.25}$$

Finally, the average symbol error probability for Case IV with BPSK modulation is given by

$$P_e = \int_0^{2\pi}\int_0^{\infty} Q\left(\sqrt{\frac{n^2E_s\gamma}{2\sigma_W^2}}\right) \frac{\exp\left[-\frac{\left(\mu_C-\sqrt{\gamma}\cos(\zeta)\right)^2}{2\sigma_C^2} - \frac{(\mu_S-\sqrt{\gamma}\sin(\zeta))^2}{2\sigma_S^2}\right]}{4\pi\sigma_C\sigma_S} d\gamma d\zeta, \tag{4.26}$$

As the inequality $\int_0^{2\pi} f(\gamma,\zeta)d\zeta \leq \int_0^{2\pi} f(\gamma,\zeta=2\pi)d\zeta$ is valid, we can write

$$P_e \leq \int_0^{\infty} Q\left(\sqrt{\frac{n^2E_s\gamma}{2\sigma_W^2}}\right) \frac{\exp\left(-\frac{\left[\mu_C-\sqrt{\gamma}\right]^2}{2\sigma_C^2} - \frac{\mu_S^2}{2\sigma_S^2}\right)}{4\pi\sigma_C\sigma_S} d\gamma. \tag{4.27}$$

Again applying the Chernoff bound, then the upper bound is given by

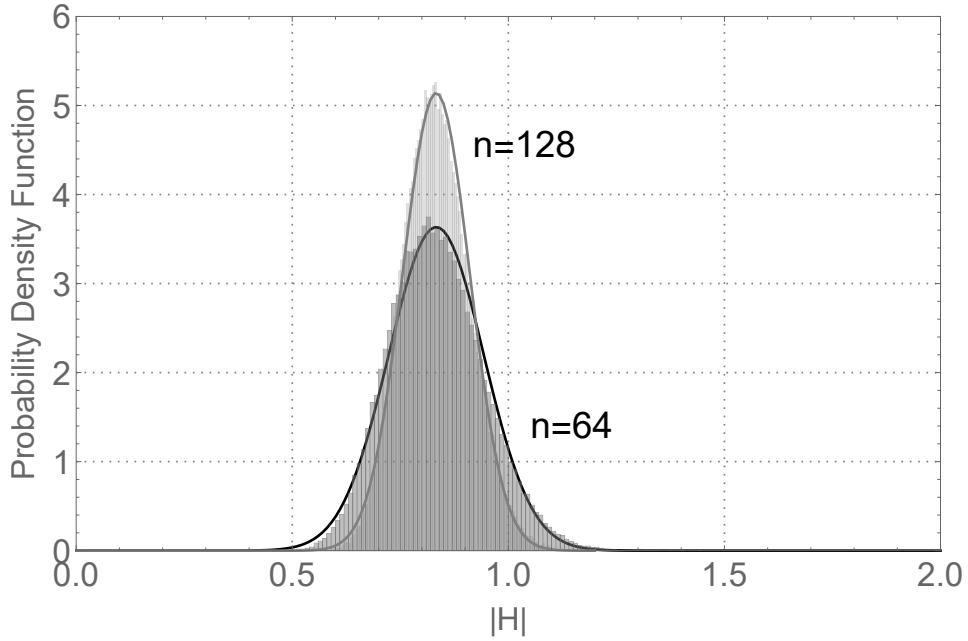
$$P_e \leq \frac{\sigma_W^3 e^{-\frac{\sigma_C^2\mu_S^2+\mu_C^2\sigma_S^2}{2\sigma_C^2\sigma_S^2}} \left(\sqrt{2\pi}\mu_S e^{\frac{\mu_S^2\sigma_W^2}{2E_sn^2\sigma_S^2+2\sigma_W^2\sigma_S^2}} \left(\operatorname{erf}\left(\frac{\mu_S}{\sqrt{2}\sqrt{n}\sigma_S^2\sqrt{\frac{E_sn}{\sigma_W^2}+\frac{1}{n\sigma_S^2}}}\right)+1\right) + 2\sqrt{n}\sigma_S^2\sqrt{\frac{E_sn}{\sigma_W^2}+\frac{1}{n\sigma_S^2}}\right)}{4\sigma_C\left(E_sn^2\sigma_S^2+\sigma_W^2\right)^{3/2}}, \tag{4.28}$$

in which erf(.) represents the error function. Also, through asymptotic analysis, it is possible to show that $\lim_{n\to\infty} P_e = 0$.

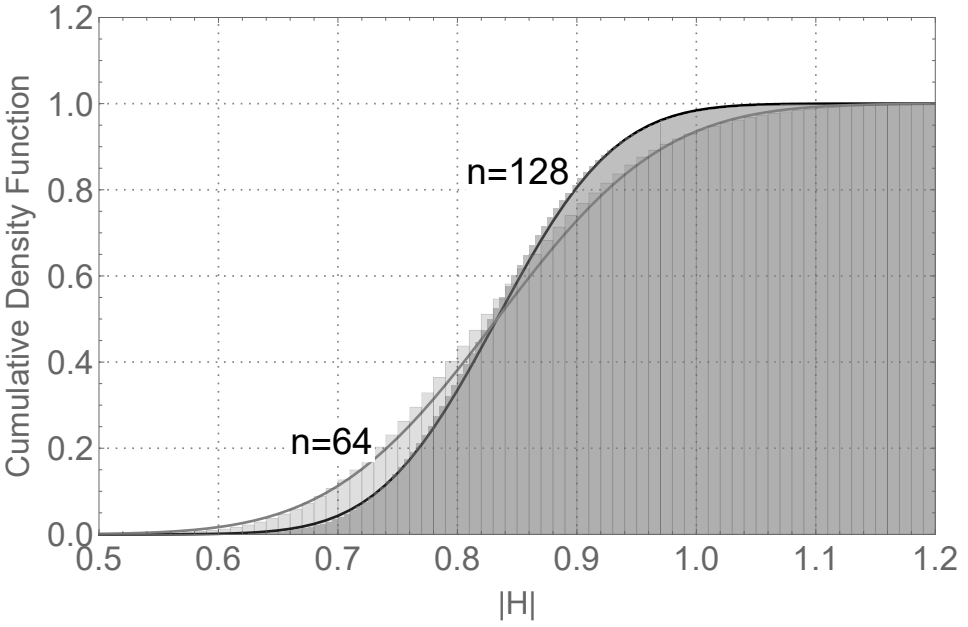
4.3 Numerical Results

In this section, we present some numerical results to validate the Monte Carlo simulations obtained from 10^5 realizations. The variables α_i and δ_i are sorted randomly at the beginning of each simulation. Except for ϕ, $\beta(\theta_i)$, and θ_i, the other parameters are the same for all cases. Note that the cases I and II are already published in the literature and are presented here for comparison only.

The probability and cumulative density function of $|H|$, given in (4.9), for Case III, is shown in Fig. 4.2 considering the following scenario: $\phi = .43\pi$, $\beta_{\min} = .2$ and $k = 1.6$ for $n = 64$ and $n = 128$.



(a)



(b)

Figure 4.2 – Probability and cumulative density function of $|H|$ for $n = 64$ and $n = 128$ for Case III.

Figure 4.3 shows the probability density function of $|H|^2$ for Case IV. The physical parameters are the same as in Figure 4.2, except $\phi = 0$ and $n \in \{16, 64\}$. In both figures, we can observe an excellent agreement between the simulations and our proposed approximation. As n increases, it is evident that the accuracy becomes better. But it is also important to mention that the resulting channel worsens significantly since $|H| \to 0$.

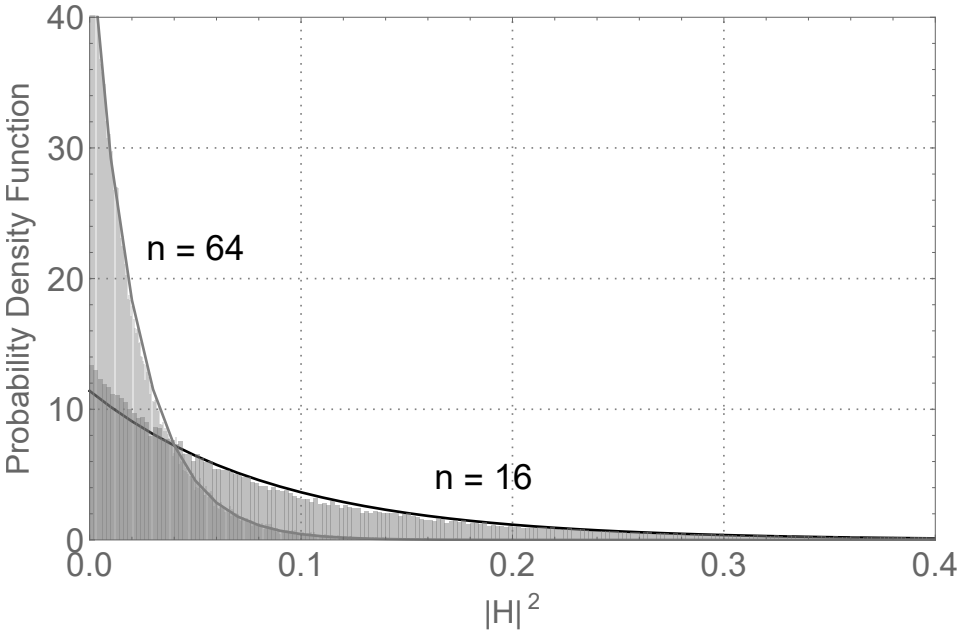


Figure 4.3 – The probability density function of $|H|^2$ for $n = 16$ and $n = 64$ for Case IV.

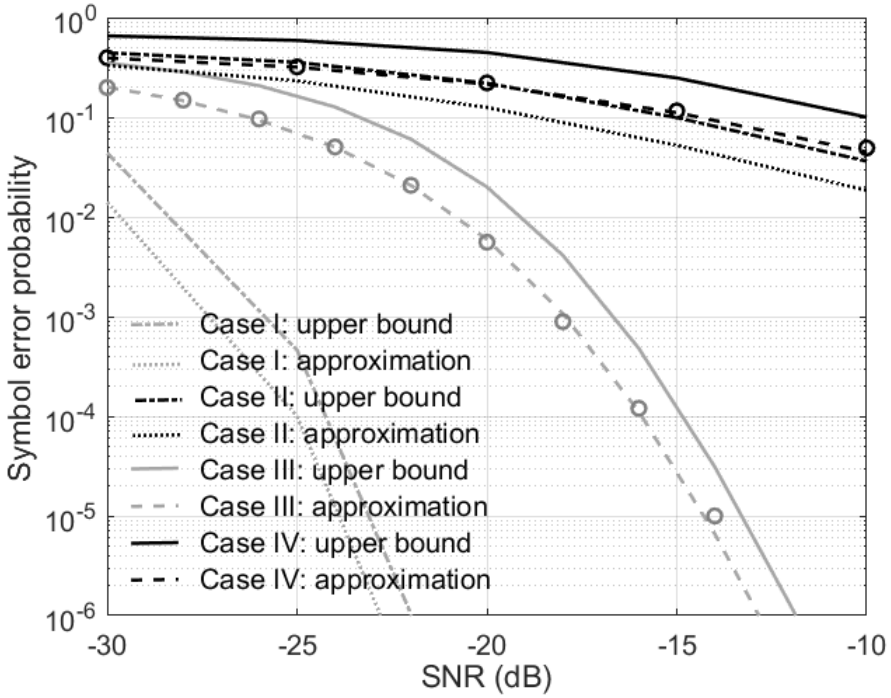


Figure 4.4 – Symbol error probability as a function of SNR. The circles represent the simulations for each case.

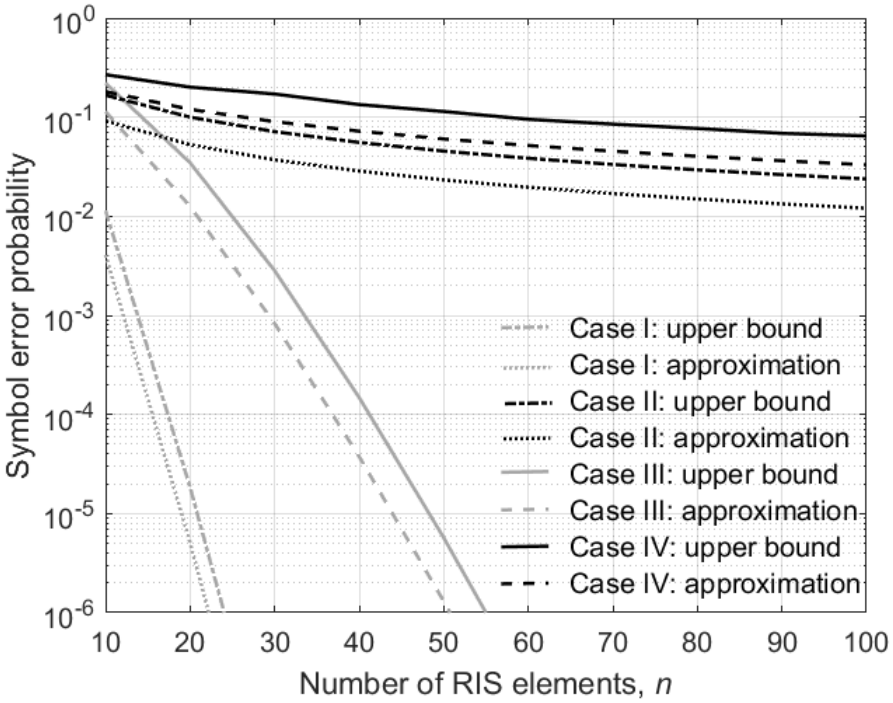


Figure 4.5 – Symbol error probability as a function of n for all cases and SNR equal to -10dB.

Figures 4.4 and 4.5 show the symbol error probability for the analytical approximation, simulation, and upper bound, as a function of SNR with $n = 64$ and n with SNR equal to 0 dB, respectively. We have considered the Cases I, II, III, and IV. The gray color represents scenarios without phase errors, while the black lines show the cases with phase errors. It is evident how close are the simulation represented by circles and the approximate curves. As expected, the SNR values are low since the resulting channel is given by the product of two others.

Also, when $\beta(\theta_i) \neq 1$ and there is phase error (Case IV), it can be seen a significant increase in the symbol error probability. Note that as the SNR increases, the gap between the cases with and without phase errors becomes even more significant. As expected, the symbol error probability decreases as the number of LIS elements increases. However, this fall is subtle for Cases II and IV.

In Figure 4.4, for a symbol error probability of 10^{-5}, the gap between Case I (no phase error and unitary reflection coefficient) and Case III (no phase error and non-unitary reflection coefficient) is around 8 dB. This number provides an idea of the impact of the reflection coefficient on the symbol error probability.

4.4 Conclusions

In this chapter, we have considered practical imperfections for the values of reflection coefficients, since they are not always unitary, as reported in the literature. Based on the

Gaussian approximation for the in-phase and quadrature components of the channel H, symbol error probability expressions were obtained for SISO scenarios with and without phase errors. These simple approximations depend only on the physical system parameters. Also, we have obtained an upper bound to the symbol error probability. All the results were validated by numerical simulations and have shown an excellent agreement. Comparisons with models in which $|\beta(\theta_i)| = 1$ (Case III and Case IV) show that practical systems present a considerable loss in performance when compared to the ideal models published in the literature (Cases I and Case II).

5

Large Intelligent Surfaces With Discrete Set of Phase-Shifts

5.1 Introduction

There is no doubt that quantization errors are inevitable when using analog-to-digital converters (ADCs). These converters bridge the analog and the digital worlds, and the lower is their resolution, the more distortions they can cause to the conversion process. Since the rounding quantization introduces error in the signal estimation stage, Hou et al. [52] propose a quantization error reduction scheme for detection based on orthogonal lattices. On the other hand, Kotera et al. [53] proves that an efficient nonlinear Viterbi-like algorithm, used as the equalization scheme, can estimate both inter-symbol interference in multi-path channel and quantization error in ADC and improve the bit error rate performance.

However, for LIS-assisted systems, little is known about the impact of quantization errors. Also called Reconfigurable Intelligent Surfaces (RIS), this technology is a strong candidate to integrate the sixth generation of cellular networks. It consists of many electromagnetic elements acting individually as scatterers capable of jointly reflecting the incident signal to the desired direction [54]. Among its advantages, we can mention the ideally passive nature that does not require any dedicated energy source. An LIS provides an inherently full-duplex transmission scheme without introducing noise, unlike relays, by not amplifying the incident signal. Besides, it can be easily installed onto facades of buildings or walls of rooms thanks to its lightweight and conformal geometry. Two strategies are possible due to the smart adjustment of the phase shifts, the reflected signals can add coherently or destructively at the receiver. The first strategy improves the received signal power, while the second one avoids interference of unwanted signals or transmitters and increases the security of the communication system [55].

Taking advantage of LIS-assisted systems' low power consumption, [56] jointly optimizes the transmit beamforming at the BS and the phase shifts at the LIS. From derived lower bounds of the transmit power concerning the number of antennas at the BS, the number of LIS elements, and the number of mobile users, they show that the transmit power at the BS is significantly lower than that of a communications system without LIS.

Although some studies optimize the reflection coefficients (i.e., amplitude and phase) of each LIS element [57], the reflection phases' high precision configuration is unfeasible since, in

practice, the number of bits dedicated to mapping a continuous number into discrete is limited. As a consequence, phase quantization errors arise. Before proposing techniques to reduce them, we first need to know them and estimate their effects as closely as possible to reality.

Badiu et al. do a preliminary analysis based on a limited number of reflectors and conclude that the performance measured from the error probability is robust against the phase errors [48]. Samith et al. [58] also consider a practical phase-shift model, but to maximize the achievable rate through the joint optimization of the transmit beamforming and the LIS reflect beamforming. On the other hand, Han et al. [59] propose an optimal phase shift design that achieves approximately the ergodic capacity and ensures that a quantizer with two bits is sufficient for a capacity degradation below 1 bit/s/Hz.

As can be seen, there are many works in the literature about the optimization of LIS-assisted systems. In this work, we deviate a little from this idea and look for more precise mathematical models in the face of possible scenarios.

In our previous work [60], we have used the Central Limit Theorem (CLT) to derive the bit error rate when there are phase estimation errors. However, it is known that the CLT is inaccurate when the number of elements in LIS is small, and the approximation error can be significant in the high Signal-to-Noise ratio (SNR) regime.

This time, we do an in-depth investigation of LIS-assisted Single-Input Single-Output (SISO) systems when there are quantization errors. Considering channels between source and destination characterized as the double (cascaded) Rayleigh fading distributions [61–64], we derive exact closed-form expressions for the spectral efficiencies, outage probabilities, and average symbol error rate (SER). Our analysis also extends to power scaling law and the power required to achieve specific capacity. Based on our excellent accuracy approach, we evaluate the system performance as the number of bits and reflectors increases. We conclude that the LIS with approximately fifty elements and four dedicated bits for phase quantization outperforms the conventional system performance without LIS. To the best of our knowledge, no similar results have been found in the literature.

Wang et al. [65] also consider a SISO LIS-assisted system and derive exact expressions for outage probability and diversity order without employing a CLT approach. However, they assume that each element of the LIS has only an one-bit phase shifter. We propose expressions for a more comprehensive scenario in which more bits are dedicated to the phase adjustment of the LIS elements.

The remainder of this chapter is organized as follows: Section 5.2 presents the adopted model and the preliminary assumptions. In Section 5.3, we derive exact closed-form expressions for some important performance metrics and evaluate the quantization error effects. Section 5.4 shows our setup and the results obtained from simulations with that. Finally, Section 5.5 summarizes the main conclusions.

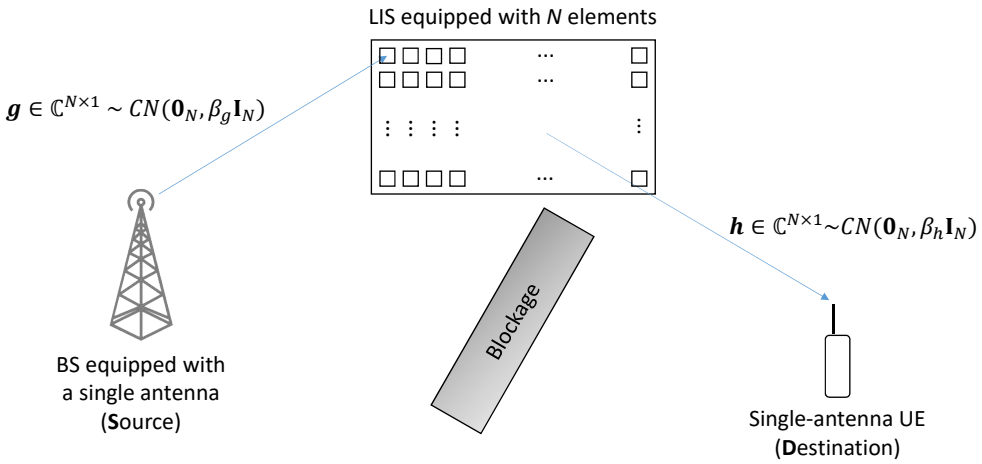


Figure 5.1 – System Model.

Notations: scalars are denoted by italic letters while vectors and matrices, by bold-face lower-case and uppercase letters, respectively. For a complex-valued vector $\mathbf{x}$, |x| denotes its Euclidean norm and diag(x) represents the diagonal matrix. The distribution of a circularly symmetric complex Gaussian (CSCG) random vector with mean x and covariance Ξ is denoted by $\mathscr{CN}(x, \Xi)$; and $\sim$ stands for "distributed as". For any general vector $\mathbf{x}$, $\mathbf{x}_i$ denote its ith element while $\mathbb{E}$ is the statistical expectation. Finally, Pr(.) represents the probability of a specific event occurring.

5.2 System Model

The system model of the adopted LIS-assisted communications scheme is shown in Fig. 5.1, where g_n and h_n represent the fading channels between the single-antenna source (S) and the nth antenna element of the LIS, and the nth antenna element of the LIS and the single-antenna destination (D), respectively. Here we assume Rayleigh fading channels, *i.e.*, $g_n \sim \mathscr{CN}(0, \beta_g)$ and $h_n \sim \mathscr{CN}(0, \beta_h)$. The parameters β_g and β_h model the shadow and geometric attenuation fading, (*i.e.*, the large-scale fading coefficients), which are assumed to be independent over the elements of LIS and change very slowly over time, being constant over several coherence-time intervals [66]. This is a reasonable assumption since the distance between devices and LIS is much larger than the distance between the LIS' elements. In this far-field regime [54], the intelligent surface is better modeled as a scatterer and the scaling law that governs the intensity of its electric field is a function of the distances' product, as proved in [67] and shown later.

We assume that the LIS is a reflect-array composed of N simple and re-configurable reflector elements connected to a controller. Additionally, we assume that the phase-shifts produced by the channels are estimated perfectly. However, the desired phases can not be accurately generated by the LIS once it has a discrete set of phases. Practical LISs have a limited number of phase shifts, *i.e.*, a discrete set of phase-shifts constrained by the number of quantization

bits (also known as phase resolution) of the LIS. The number of quantization bits is denoted by b. Therefore, the set of phase shifts produced by each one of the elements of the LIS is defined as

$$\phi_n = \left\{0, \frac{2\pi}{2^b}, \frac{4\pi}{2^b}, \cdots, \frac{2\pi(2^b-1)}{2^b}\right\}. \tag{5.1}$$

Therefore, we model the deviation from the correct/ideal phase-shift as a phase-noise, δ_n, which is uniformly distributed in the range $[-\pi/Q, \pi/Q]$, where $Q = 2^b$ is the number of discrete phases the LIS can generate [68] dictated by the hardware complexity and power consumption of LIS.

5.3 Intelligent Transmission Through LIS

In slowly varying flat fading channels, the signal received at the destination after being reflected through an LIS composed of N passive elements can be written as

$$y = \sqrt{\rho}\left[\sum_{n=1}^{N} g_n e^{-j\phi_n} h_n\right] s + w, \tag{5.2}$$

where ρ is the average SNR, ϕ_n is the adjustable phase-shift produced by the nth LIS reflector, s is the modulation data symbol with zero mean, $\mathbb{E}[|s|^2] = 1$, and $w \sim \mathscr{CN}(0,1)$ is the additive white Gaussian noise (AWGN) term.

Then, (5.2) can be re-written in the matrix-form as

$$y = \sqrt{\rho}\mathbf{h}^T \Phi \mathbf{g} s + w, \tag{5.3}$$

where $\mathbf{g} = [g_1, \cdots, g_N]^T$ and $\mathbf{h} = [h_1, \cdots, h_N]^T$ are the channel coefficient vectors between the BS and the RIS and between the RIS and the terminal, respectively, while $\Phi = \text{diag}\left(\left[e^{-j\phi_1}, \cdots, e^{-j\phi_N}\right]\right)$ is the diagonal matrix containing the phase-shifts applied by the elements of the LIS.

It can be noticed that (5.3) is similar to the equation of conventional MIMO systems employing precoding/beamforming for transmission. However, differently from those systems, where precoding/beamforming is carried out at the transmitter side, here it is carried out over the transmission medium (*i.e.*, the environment) [69].

The complex channels can be written in polar representation (*i.e.*, with magnitude and phase) as $h_n = \alpha_n e^{j\theta_n}$ and $g_n = \xi_n e^{j\psi_n}$, therefore, (5.2) can be re-written as

$$\begin{aligned} y &= \sqrt{\rho}\left[\sum_{n=1}^{N} \alpha_n \xi_n e^{j(\theta_n+\psi_n-\phi_n)}\right] s + w \\ &= \sqrt{\rho}\left[\sum_{n=1}^{N} \alpha_n \xi_n e^{j\delta_n}\right] s + w, \end{aligned} \tag{5.4}$$

where the second line is obtained from the assumption that the LIS only generates discrete phases and consequently, there is a phase-noise, $\delta_n = \theta_n + \psi_n - \phi_n$.

Considering the phase-noise, then the instantaneous SNR at the destination is given by

$$\gamma = \rho \left| \sum_{n=1}^{N} \alpha_n \xi_n e^{j\delta_n} \right|^2 . \tag{5.5}$$

Note that the instantaneous SNR is maximized when $\delta_n = 0$, *i.e.*, the channels are perfectly estimated, and the LIS can accurately generate the phases induced by the channels (meaning that $Q \to \infty$) [50].

Lemma 1 *From empirical comparisons between the normalised histogram of the random variable given by*

$$r = \sqrt{\rho} \left| \sum_{n=1}^{N} \alpha_n \xi_n e^{j\delta_n} \right| = \sqrt{\rho} \left| \sum_{n=1}^{N} |g_n||h_n| e^{j\delta_n} \right|, \tag{5.6}$$

and the theoretical PDF of a Gamma random variable, it is possible to say that the PDF of r can be accurately approximated by the Gamma PDF with shape and scale parameters given by κ and θ, respectively as

$$\kappa = \frac{-\left(\mathbb{E}\left[\gamma^2\right] - 5\mathbb{E}\left[\gamma\right]^2\right) + \sqrt{\mathbb{E}\left[\gamma^2\right]^2 - 34\mathbb{E}\left[\gamma^2\right]\mathbb{E}\left[\gamma\right]^2 + 49\mathbb{E}\left[\gamma\right]^4}}{2\left(\mathbb{E}\left[\gamma^2\right] - \mathbb{E}\left[\gamma\right]^2\right)} > 0, \tag{5.7}$$

$$\theta = \sqrt{\frac{-\sqrt{\mathbb{E}\left[\gamma^2\right]^2 + 14\mathbb{E}\left[\gamma^2\right]\mathbb{E}\left[\gamma\right]^2 + \mathbb{E}\left[\gamma\right]^4} + 2\mathbb{E}\left[\gamma^2\right] + 2\mathbb{E}\left[\gamma\right]^2}{6\mathbb{E}\left[\gamma\right]}} > 0, \tag{5.8}$$

where $\mathbb{E}\left[\gamma\right]$ and $\mathbb{E}\left[\gamma^2\right]$ are given by (5.11) *and* (5.12)*, respectively.*

Some examples of this comparison are shown in Section 5.4. The parameters κ and θ are found following the rationale presented in Appendix B.1. Therefore, the PDF of γ can be found following the standard transformation of random variables, $\gamma = r^2$, and is defined as

$$f_\gamma(\gamma) = \frac{1}{2\Gamma(\kappa)\theta^\kappa} \gamma^{\left(\frac{\kappa-2}{2}\right)} e^{-\frac{\sqrt{\gamma}}{\theta}}, \gamma \geq 0. \tag{5.9}$$

In its turn, the cumulative distribution function (CDF) of the SNR random variable, γ, is defined as

$$F_\gamma(\gamma) = \int_0^\gamma f_\gamma(x)dx = 1 - \frac{\Gamma\left(\kappa, \frac{\sqrt{\gamma}}{\theta}\right)}{\Gamma(\kappa)}, \gamma \geq 0, \tag{5.10}$$

where $\Gamma(.)$ is the Euler gamma function while $\Gamma(.,.)$ is the upper incomplete gamma function. The integral result is obtained by directly applying (Eq. 2.33.10, [70]).

Remark 1 *When $Q \to \infty$, i.e., the LIS is able to generate any phase-shift, the phase-noise is zero, $\delta_n = 0, \forall n$, and consequently,* (5.11) *and* (5.12) *can be simplified to* (5.13) *and* (5.14)*, respectively, and whose derivations are presented in Appendix B.2.*

$$\mathbb{E}\left[\gamma\right] = \mathbb{E}\left[r^2\right] = \rho\beta_g\beta_h\mathscr{A}_1 = \rho\beta_g\beta_h N\left[1 + \frac{1}{16}(N-1)Q^2\sin^2\left(\frac{\pi}{Q}\right)\right]. \tag{5.11}$$

$$\begin{aligned}\mathbb{E}\left[\gamma^2\right] &= \mathbb{E}\left[r^4\right] = (\rho\beta_g\beta_h)^2\mathscr{A}_2 \\ &= (\rho\beta_g\beta_h)^2\frac{N}{256}\left\{512(N+1) + \frac{32(N-1)Q^2}{\pi^2} + \frac{(N-1)Q^2\left[\pi\sin^2\left(\frac{\pi}{Q}\right)\left((N-2)Q\left(\pi(N-3)Q\sin^2\left(\frac{\pi}{Q}\right) + 16\sin\left(\frac{2\pi}{Q}\right)\right) + 16\pi(4N+1)\right) - 32\cos\left(\frac{4\pi}{Q}\right)\right]}{\pi^2}\right\}.\end{aligned} \tag{5.12}$$

$$\lim_{Q\to\infty}\mathbb{E}\left[\gamma\right] = N\rho\beta_g\beta_h\left[1 + \frac{(N-1)\pi^2}{16}\right]. \tag{5.13}$$

$$\lim_{Q\to\infty}\mathbb{E}\left[\gamma^2\right] = \frac{N\rho^2\left(\beta_g\beta_h\right)^2}{256}\left[256 + 768N + \pi^4(N-3)(N-2)(N-1) + 48\pi^2(2N-1)(N-1)\right]. \tag{5.14}$$

5.3.1 Exact Ergodic Spectral Efficiency

The ergodic spectral efficiency of the LIS-assisted system is defined as

$$\begin{aligned}C &= \mathbb{E}\left[\log_2\left(1+\gamma\right)\right] \\ &= \int_0^\infty \log_2\left(1+\gamma\right) f_\gamma(\gamma)d\gamma.\end{aligned} \tag{5.15}$$

Solving (5.15) through an integral solver [37], we find the exact closed-form expression for the ergodic capacity given by (5.16), where ${}_pF_q\left(a_1,\cdots,a_p;b_1,\cdots,b_q;z\right)$ is the generalized hypergeometric function [71] and $\Psi^{(n)}(z)$ is the nth derivative of the digamma function, also known as the polygamma function [72].

$$\begin{aligned}C = &\frac{2\,{}_2F_3\left(1,1;2,\frac{3}{2}-\frac{\kappa}{2},2-\frac{\kappa}{2};-\frac{1}{4\theta^2}\right)}{\theta^2(\kappa-1)(\kappa-2)\log(4)} - \frac{2\pi\sec\left(\frac{\pi\kappa}{2}\right)\,{}_1F_2\left(\frac{\kappa}{2}+\frac{1}{2};\frac{3}{2},\frac{\kappa}{2}+\frac{3}{2};-\frac{1}{4\theta^2}\right)}{(\kappa+1)\theta^{\kappa+1}\Gamma(\kappa)\log(4)} \\ &+ \frac{2\pi\csc\left(\frac{\pi\kappa}{2}\right)\,{}_1F_2\left(\frac{\kappa}{2};\frac{1}{2},\frac{\kappa}{2}+1;-\frac{1}{4\theta^2}\right)}{\kappa\theta^\kappa\Gamma(\kappa)\log(4)} + \frac{4(\log(\theta)+\psi^{(0)}(\kappa))}{\log(4)}.\end{aligned} \tag{5.16}$$

Remark 2 *In high SNR regime, the ergodic spectral efficiency in* (5.16) *can be approximated as*

$$C_{\text{high-SNR}} \approx \frac{2}{\theta^2(\kappa-1)(\kappa-2)\log(4)} - \frac{2\pi\sec\left(\frac{\pi\kappa}{2}\right)}{(\kappa+1)\theta^{\kappa+1}\Gamma(\kappa)\log(4)} + \frac{2\pi\csc\left(\frac{\pi\kappa}{2}\right)}{\kappa\theta^\kappa\Gamma(\kappa)\log(4)} + \frac{4(\log(\theta)+\psi^{(0)}(\kappa))}{\log(4)}. \tag{5.17}$$

Remark 3 *When $\rho\to\infty$, then* (5.16) *becomes* (5.18).

$$\lim_{\rho\to\infty} C = \frac{4\log(\theta)}{\log(4)}. \tag{5.18}$$

The proofs for Remarks 2 and 3 are presented in Appendices B.3 and B.4, respectively.

Remark 4 *In high SNR and N regimes, the ergodic spectral efficiency can be approximated as*

$$C_{\textit{high-SNR, N}} \approx \frac{2}{\theta^2(\kappa-1)(\kappa-2)\log(4)} + \frac{4(\log(\theta)+\psi^{(0)}(\kappa))}{\log(4)}. \tag{5.19}$$

The proof for Remark 4 is presented in Appendix B.5.

Other alternative to find the expectation in (5.15) is using the PDF of the random variable given by $C_{\text{inst.}} = \log_2(1+\gamma)$, i.e., the instantaneous spectral efficiency, which can be found after applying standard transformation of random variables to (5.9) giving rise to

$$f_{C_{\text{inst.}}}(c) = \frac{\log(2)}{\Gamma(\kappa)\theta^{\kappa}} 2^{c-1}(2^c-1)^{\left(\frac{\kappa-2}{2}\right)} e^{-\frac{\sqrt{2^c-1}}{\theta}}, c \geq 0. \tag{5.20}$$

Then, the CDF of the instantaneous spectral capacity random variable is expressed by

$$F_{C_{\text{inst.}}}(c) = \int_0^c f_{C_{\text{inst.}}}(x)\,dx = 1 - \frac{\Gamma\left(\kappa, \frac{\sqrt{2^c-1}}{\theta}\right)}{\Gamma(\kappa)}, \tag{5.21}$$

whose integral is also found by directly applying (Eq. 2.33.10, [70]).

5.3.2 Upper and Lower-bounds for the Ergodic Spectral Efficiency

As it can be seen, (5.16) is quite complex. Therefore, here we aim at finding simpler but yet tight bounds for the ergodic spectral efficiency of the LIS-assisted system. According to Jensen's inequality [66], it holds that

$$\mathbb{E}\left[\log_2(1+\gamma)\right] \leq \log_2\left(1+\mathbb{E}\left[\gamma\right]\right). \tag{5.22}$$

Then, by using $\mathbb{E}[\gamma]$ given by (5.11), a possible upper-bound for the ergodic capacity of the LIS-assisted system can be given by (5.23).

$$C \leq C_{\text{upper}} = \log_2\left(1 + N\rho\beta_g\beta_h\left[1+\frac{1}{16}(N-1)Q^2\sin^2\left(\frac{\pi}{Q}\right)\right]\right). \tag{5.23}$$

As it is tight for high SNR scenarios, it can be assumed as a good approximation. On the other hand, again according to Jensen's inequality [66], it holds that

$$\mathbb{E}\left[\log_2(1+\gamma)\right] \geq \log_2\left(1+\left[\mathbb{E}\left[\frac{1}{\gamma}\right]\right]^{-1}\right). \tag{5.24}$$

Consequently, by using a tight approximation of $\mathbb{E}[1/\gamma]$ (see Appendix B.6), a lower-bound for the ergodic capacity of the LIS-assisted system can be derived and given as

$$\begin{aligned} C \geq C_{\text{lower}} &\approx \log_2\left(1+\frac{\mathbb{E}[\gamma]^3}{\mathbb{E}[\gamma^2]}\right) \\ &= \log_2\left(1+\frac{256N^2\rho\beta_g\beta_h\left(\frac{1}{16}(N-1)Q^2\sin^2\left(\frac{\pi}{Q}\right)+1\right)^3}{\frac{32(N-1)Q^2}{\pi^2}+\frac{(N-1)Q^2\left(\pi\sin^2\left(\frac{\pi}{Q}\right)\left((N-2)Q\left(\pi(N-3)Q\sin^2\left(\frac{\pi}{Q}\right)+16\sin\left(\frac{2\pi}{Q}\right)\right)+16\pi(4N+1)\right)-32\cos\left(\frac{4\pi}{Q}\right)\right)}{\pi^2}+512(N+1)}\right). \end{aligned} \tag{5.25}$$

Like the SNR, the spectral efficiency is also maximized when $Q \to \infty$, meaning that the LIS has infinite phase-shift precision and can generate any phase-shift. In this case, the maximum ergodic spectral efficiency with the upper and lower bounds are given, respectively, by

$$\begin{aligned} C_{\text{upper}}^{\text{max.}} &= \lim_{Q\to\infty} C_{\text{upper}} \\ &= \lim_{Q\to\infty} \log_2\left(1 + N\rho\beta_g\beta_h\left[1 + \frac{1}{16}(N-1)Q^2\sin^2\left(\frac{\pi}{Q}\right)\right]\right) \\ &= \log_2\left(1 + N\rho\beta_g\beta_h\left[1 + \frac{(N-1)\pi^2}{16}\right]\right). \end{aligned} \tag{5.26}$$

and

$$\begin{aligned} C_{\text{lower}}^{\text{max.}} &= \lim_{Q\to\infty} C_{\text{lower}} \\ &= \log_2\left(1 + \frac{N^2\rho\beta_g\beta_h\left(\pi^2(N-1)+16\right)^3}{16\left((N-1)\left(\pi^4\left(N^2-5N+6\right)+48\pi^2(2N-1)+256\right)+512(N+1)\right)}\right). \end{aligned} \tag{5.27}$$

5.3.3 Impact of Bit Quantization in the Spectral Efficiency

In practical communication systems, the set of phase-shifts is limited by the number of quantization bits of the LIS, influencing the achieved spectral efficiency directly. Therefore, in this section, we propose a criterion for selecting the number of quantization levels $Q = 2^b$ so that the ergodic spectral efficiency is optimized up to a specific spectral degradation in bits/s/Hz. In order to quantify this degradation, we define the error ϵ brought about by a limited number of phase-shifts as

$$C_{\text{upper}}^{\text{max.}} - C_{\text{upper}} \leq \epsilon. \tag{5.28}$$

Remark 5 *From* (5.28)*, we see that when the number of LIS elements tends to* ∞*, then the ergodic spectral efficiency degradation,* ϵ*, becomes*

$$\lim_{N\to\infty} \epsilon = \log_2\left(\frac{\pi^2}{Q^2\sin^2\left(\pi/Q\right)}\right) \textit{ bits/s/Hz.} \tag{5.29}$$

Remark 6 *From* (5.28)*, we see that when* $\rho \to \infty$*, then the ergodic spectral efficiency degradation,* ϵ*, is given by*

$$\lim_{\rho\to\infty} \epsilon = \log_2\left(\frac{16 + (N-1)\pi^2}{16 + (N-1)Q^2\sin^2\left(\pi/Q\right)}\right) \textit{ bits/s/Hz.} \tag{5.30}$$

Proposition 1 *In order to guarantee an suitable ergodic spectral efficiency degradation of* ϵ *bits/s/Hz compared to an LIS with full-resolution phase-shift, the number of quantization levels,* Q*, of the LIS should satisfy*

$$Q\sin\left(\pi/Q\right) \geq \sqrt{\frac{16\left(2^{-\epsilon}-1\right)}{N\rho\beta_g\beta_h(N-1)} + \frac{16\left(2^{-\epsilon}-1\right)}{N-1} + \pi^2 2^{-\epsilon}}. \tag{5.31}$$

Remark 7 *From* (5.31)*, we see that when* $N \to \infty$*, the number of quantization levels,* Q*, should satisfy*

$$\lim_{N\to\infty} Q \sin(\pi/Q) \geq \sqrt{2^{-\epsilon}}\pi. \tag{5.32}$$

Remark 8 *From* (5.31) *we see that when* $\epsilon \to \infty$*, then the number of quantization levels,* Q*, should satisfy*

$$\lim_{\epsilon\to\infty} Q \sin(\pi/Q) \leq \pi. \tag{5.33}$$

After analysing Remark 8, we notice that the first term in (5.33) is equal to π only when $Q \to \infty$. Therefore, in order to have no degradation at all, an infinite number of quantization levels is necessary, which demonstrates the correctness of Remark 8.

Summing up, these results can be used to select the precision necessary for an LIS-assisted system to achieve a pre-defined and acceptable degradation in its ergodic spectral efficiency.

5.3.4 Outage Probability

Based on the knowledge of the approximate PDF of the instantaneous spectral efficiency given by (5.20), it is possible to find its cumulative density function (CDF) and derive analytical expressions for the outage probability. The outage probability is defined as the probability that the achieved instantaneous spectral efficiency falls below a given threshold $C_{\text{out.}}$ and can be written as

$$\begin{aligned} P_{\text{out.}} &= \Pr\{C_{\text{inst.}} < C_{\text{out.}}\} \\ &= \int_0^{C_{\text{out.}}} f_{C_{\text{inst.}}}(x)\, dx \\ &= 1 - \frac{\Gamma\left(\kappa, \frac{\sqrt{2^{C_{\text{out.}}}-1}}{\theta}\right)}{\Gamma(\kappa)}, \end{aligned} \tag{5.34}$$

whose proof is provided in Appendix B.7.

Besides that way, the outage probability can also be defined with regard to the instantaneous SNR. In this case,it is the probability that the instantaneous SNR falls below a given SNR threshold $\gamma_{\text{out.}}$. So, the outage probability is given by

$$\begin{aligned} P_{\text{out.}} &= \Pr\{\gamma < \gamma_{\text{out.}}\} \\ &= \int_0^{\gamma_{\text{out.}}} f_{\gamma}(x)\, dx \\ &= \frac{1}{\theta^{\kappa/2}} \left[1 - \frac{\Gamma\left(\kappa, \frac{\sqrt{\gamma_{\text{out.}}}}{\theta}\right)}{\Gamma(\kappa)}\right], \end{aligned} \tag{5.35}$$

and found by using (B.48) in Appendix B.7. It can also be expressed as

$$P_{\text{out.}} = \frac{\gamma^{\frac{\kappa}{2}}}{\kappa\theta^{\frac{3\kappa}{2}}} {}_1F_1\left(\kappa, \kappa+1, -\frac{\sqrt{\gamma}}{\theta}\right). \tag{5.36}$$

Remark 9 *In high SNR regime, the outage probability can be approximated as*

$$P_{out.}^{high\text{-}SNR} = \frac{\gamma^{\frac{\kappa}{2}}}{\kappa\theta^{\frac{3\kappa}{2}}}. \tag{5.37}$$

The proofs of (5.36) and (5.37) are provided in Appendix B.8.

5.3.5 Average Symbol Error Rate

According to [51], the average Symbol Error Rate (SER) is defined as the expectation of conditional error probability, $P_{e|\gamma}$, given the distribution of the SNR, γ. For a wide variety of modulation schemes, $P_{e|\gamma}$ is defined as $P_{e|\gamma} = \bar{a}\mathscr{Q}\left(\sqrt{\bar{b}\gamma}\right)$, where $\bar{a}$ and $\bar{b}$ are constant modulation dependent parameters and $\mathscr{Q}$ is the Gaussian $\mathscr{Q}$-function defined as $\int_x^\infty e^{-t^2/2}/\sqrt{2\pi}dt$ [51]. Therefore, the average SER is derived as

$$\mathbb{E}\left[\bar{a}\mathscr{Q}\left(\sqrt{\bar{b}\gamma}\right)\right] = \bar{a}\int_0^\infty \mathscr{Q}\left(\sqrt{\bar{b}\gamma}\right) f_\gamma(\gamma)d\gamma, \tag{5.38}$$

and can be analytically expressed by

$$P_e = \mathbb{E}\left[\bar{a}\mathscr{Q}\left(\sqrt{\bar{b}\gamma}\right)\right] = \bar{a}2^{-\frac{\kappa}{2}-1}\bar{b}^{-\frac{\kappa}{2}}\theta^{-\kappa}\times$$
$$\left(\frac{{}_2F_2\left(\frac{\kappa}{2}+\frac{1}{2}, \frac{\kappa}{2}; \frac{1}{2}, \frac{\kappa}{2}+1; \frac{1}{2\bar{b}\theta^2}\right)}{\Gamma\left(\frac{\kappa}{2}+1\right)} - \frac{\kappa\,{}_2F_2\left(\frac{\kappa}{2}+\frac{1}{2}, \frac{\kappa}{2}+1; \frac{3}{2}, \frac{\kappa}{2}+\frac{3}{2}; \frac{1}{2\bar{b}\theta^2}\right)}{\sqrt{2\bar{b}}\theta\Gamma\left(\frac{\kappa+3}{2}\right)}\right), \tag{5.39}$$

whose proof is provided in Appendix B.9. Note that in (5.39), $\bar{a}$ and $\bar{b}$ are constants that depend on the modulation scheme. For instance, the average SER of the binary phase shift keying (BPSK) modulation is obtained when $\bar{a} = 1$ and $\bar{b} = 2$, while that for the M-ary Pulse Amplitude Modulation (M-PAM), $\bar{a} = 2(M-1)/M$ and $\bar{b} = 6/(M^2-1)$. In the same way, $\bar{a} = \bar{b} = 2$ are applied for the average SER of the quadrature phase shift keying (QPSK) modulation. Finally, $\bar{a} = 2$ and $\bar{b} = 2\sin^2(\pi/M)$ for M-ary phase shift keying (M-PSK) modulation, while $\bar{a} = 4(1-1/\sqrt{M})$ and $\bar{b} = 3/(M-1)$ for the average SER of the M-ary quadrature amplitude modulation (M-QAM), when $M > 4$.

Remark 10 *In high SNR regime, the average SER can be approximated as*

$$P_e^{high\text{-}SNR} \approx \bar{a}2^{-\frac{\kappa}{2}-1}\bar{b}^{-\frac{\kappa}{2}}\theta^{-\kappa}\left(\frac{1}{\Gamma\left(\frac{\kappa}{2}+1\right)} - \frac{\kappa}{\sqrt{2\bar{b}}\theta\Gamma\left(\frac{\kappa+3}{2}\right)}\right), \tag{5.40}$$

whose proof is provided in Appendix B.10.

After analyzing (5.40), it is possible to observe that the first term inside the parentheses is the dominant one. Otherwise, the average SER would be a negative number since a, b, and θ are values greater than zero. This direct insight results in the following remark.

Remark 11 *The average SER decreases when κ and/or b increases and when a and/or θ decreases.*

As shown in Section 5.4, this remark demonstrates that the average SER decreases as the transmission power, ρ, and/or the number of reflecting elements, N, increases. On the other hand, the average SER increases as the modulation order increases.

5.3.6 Diversity Order

The diversity order is a fundamental parameter of diversity-based systems. It measures the number of independent paths over which the data is received. The diversity order, D, is formally defined as the negative slope of the average SER versus the average SNR curve in a log-log scale, and calculated as by [73]

$$D = \lim_{\rho - \infty} -\frac{\log P_e}{\log \rho}. \tag{5.41}$$

From the definition above, we can see that the diversity order is a high-SNR concept.

Remark 12 *The diversity order of the LIS-assisted system is obtained as*

$$D = \frac{5\mathscr{A}_1^2 + \sqrt{49\mathscr{A}_1^4 - 34\mathscr{A}_1^2\mathscr{A}_2 + \mathscr{A}_2^2} - \mathscr{A}_2}{4\left(\mathscr{A}_2 - \mathscr{A}_1^2\right)}. \tag{5.42}$$

The parameters and proof of (5.42) are detailed in Appendix B.11. From them, we realize that the diversity order increases as N.

Remark 13 *Despite both source and destination being equipped with a single antenna, the achievable diversity order grows with the number of LIS reflecting elements. It is worth noting that each reflecting element modifies the incident waves' phases to add at the destination coherently. A direct SISO path between source and destination would only allow for a unitary diversity order. Once diversity gains can only be obtained by employing multiple antennas at transmission and/or receiving sides. However, LIS employment provides a substantial diversity order to the communication system just by adding passive reflecting elements with adjustable phases to the system.*

5.3.7 Power-scaling law

This subsection analyses the power-scaling law of the ergodic spectral efficiency regarding the number of reflecting elements in an LIS-assisted system in which $N \to \infty$.

If N grows without limit and we consider that the transmit power, ρ, can be scaled down with N^2 according $\rho = P/N^2$ and P is fixed, then (5.23) and (5.26) become, respectively

$$\begin{aligned} C_{\text{upper}} &= \log_2\left(1 + N\frac{P}{N^2}\beta_g\beta_h\left[1 + \frac{(N-1)Q^2\sin^2\left(\frac{\pi}{Q}\right)}{16}\right]\right) \\ &\to \frac{P\beta_g\beta_h Q^2\sin^2\left(\frac{\pi}{Q}\right)}{16}, N \to \infty \end{aligned} \tag{5.43}$$

and

$$\begin{aligned} C_{\text{upper}}^{\text{max.}} &= \log_2\left(1 + N\frac{P}{N^2}\beta_g\beta_h\left[1 + \frac{(N-1)\pi^2}{16}\right]\right) \\ &\to \frac{P\beta_g\beta_h\pi^2}{16}, N \to \infty. \end{aligned} \tag{5.44}$$

These results confirm that with many reflecting elements and perfect channel information, the transmit power can be reduced proportionally to $1/N^2$ without compromising the spectral efficiency.

Remark 14 *From* (5.23) *and* (5.43)*, it is possible to see that if we decrease the transmit power proportionally to* $1/N^\alpha$*, with* $\alpha > 2$*, then the SNR goes to zero as* $N \to \infty$*. When* $\alpha < 2$*, the SNR grows without bound as* $N \to \infty$*. This means that* $1/N^2$ *(i.e.,* $\alpha = 2$*) is the fastest rate at which we can decrease the transmit power and still maintain a fixed rate.*

The Remark 14 shows that as N grows without bound, the transmit power can be reduced proportionally to $1/N^2$. The transmit power reduction is significant mainly to power-constrained devices such as IoT devices [74, 75].

5.4 Simulation Results

This section presents some numerical results to validate the derived expressions against Monte Carlo simulations obtained from 10^6 realizations. The setup in Figure 5.2 shows the geometric placement adopted for the BS, LIS and UE, where r_g and r_h are the distances between source (i.e., the BS) and LIS, and between LIS and destination (i.e., the UE), respectively. Both of them are set to 25 m.

We assume that the large-scale fading coefficients are modeled as $\beta_g = z_g/(r_g)^\nu$ and $\beta_h = z_h/(r_h)^\nu$, in which z_g and z_h are log-normal random variables with standard deviation σ_{shadow}, while r_g is the distance between the source and the LIS and r_h is the distance between the LIS and the destination. ν is the path-loss exponent. For all simulation results, we adopt the typical suburban area parameters $\sigma_{\text{shadow}} = 8$ dB and $\nu = 3.67$.

Figure 5.3 presents some comparisons of the normalized histogram of the random variable given by the instantaneous SNR (see (5.5)) against the theoretical PDF given by (5.9). Note

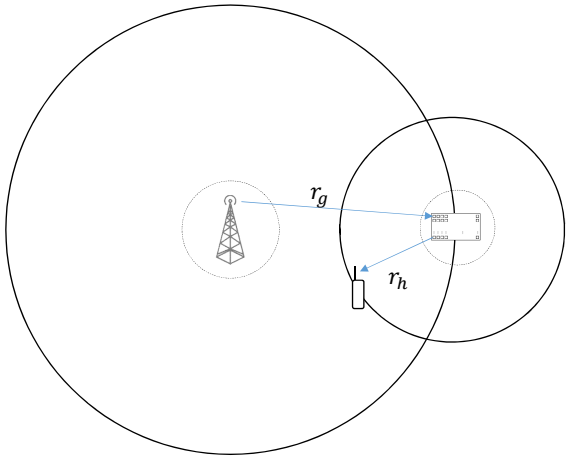


Figure 5.2 – Adopted setup.

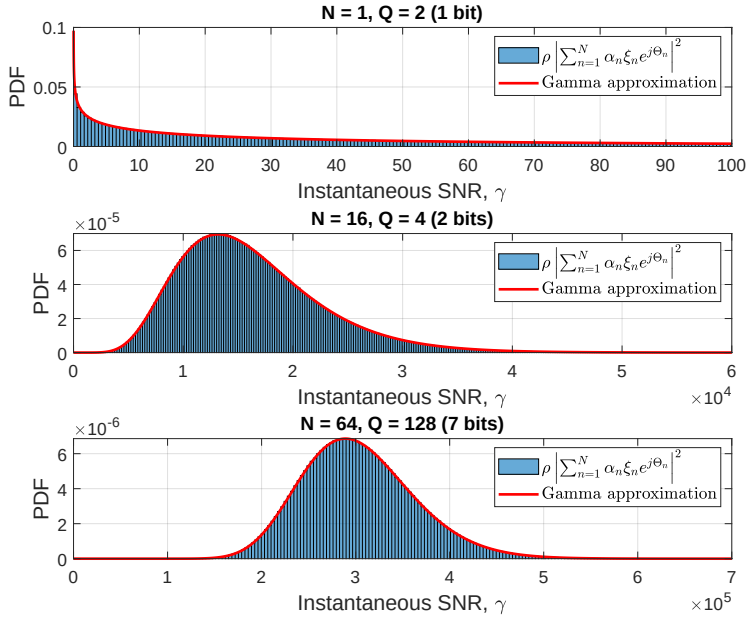


Figure 5.3 – Comparison of the approximated PDF for the instantaneous sum-capacity.

that the x axis is given in dB in the first plot, but linear scale in the others. As can be noticed, even for a small number of reflecting elements and quantization bits, the approximation is quite tight.

Figure 5.4 shows the Kullback-Leibler Divergence [76] between the approximated SNR PDF and the real distribution of the SNR random variable over the variation of the number of quantization bits and for several values of LIS elements, N. In general, this is the most known technique to evaluate an approximation in statistics. As can be seen, from $b = 2$ bits onward, the divergence remains constant regardless of the number of elements. Additionally, the figure also shows that as the number of elements increases, the divergence decreases. These results

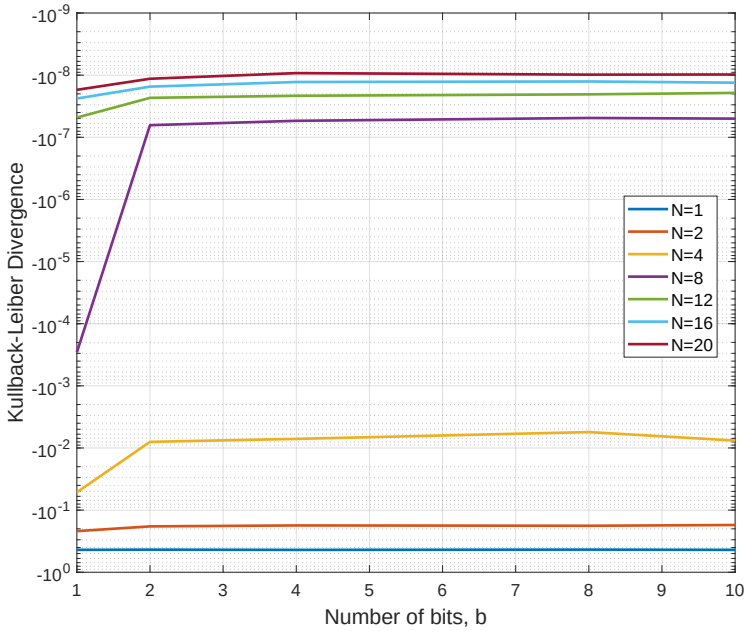


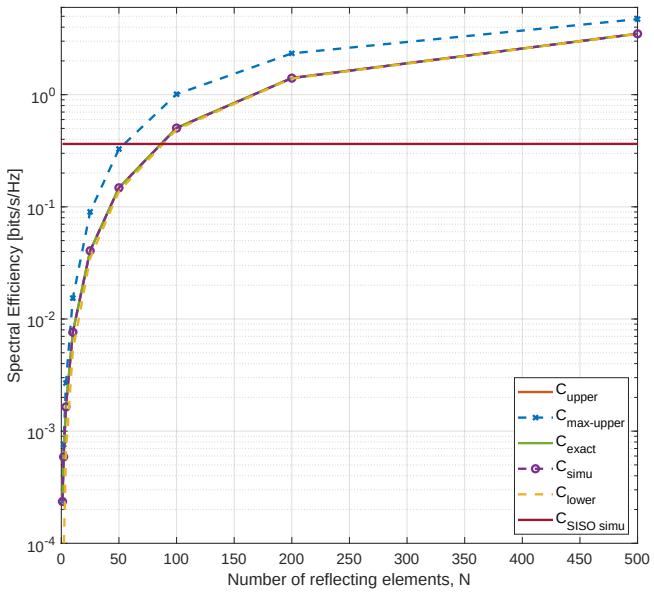
Figure 5.4 – Kullback-Leibler divergence between the approximated SNR PDF and the real distribution.

reveal that only the number of LIS reflecting elements can take the approximated PDF closer to the real one and that the number of quantization bits has a minimal impact on it. It is aligned to the theory since an inspection of (5.6) reveals that only the number of reflecting elements impact the summation in that equation.

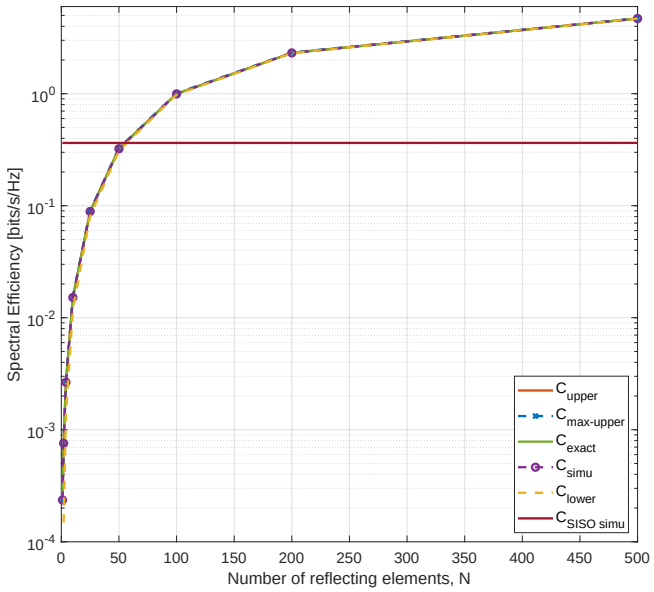
From Figure 5.5, which shows the spectral efficiency as a function of N for $b \in \{1, 4, 10\}$, we can see that the accuracy of the approximation becomes better not only as N increases but also when more bits are dedicated to phase quantization. For comparison, we also present the simulated capacity curve of a SISO system without the assistance of an LIS. When $b = 1$, the aid provided by an LIS becomes advantageous for $N > 80$. Otherwise, When $b > 1$, an LIS with $N > 50$ is enough for the LIS-aided system's behavior to outperform that of the system without an LIS.

We also verify the performance degradation when b varies. As shown in Figure 5.6, the spectral efficiency decreases when b is small. This is evident, especially for $b = 1$ and $b = 2$. Therefore, the performance difference between the spectral efficiency obtained when using perfect phase shifts and that of using quantized phase shifts decreases as b increases. Moreover, the degradation also tends to decrease as more reflective elements are added to the LIS. That is, for an LIS with many elements, a few bits (as few as 4 bits) are sufficient for quantization with a negligible performance degradation, which is in accordance (5.26).

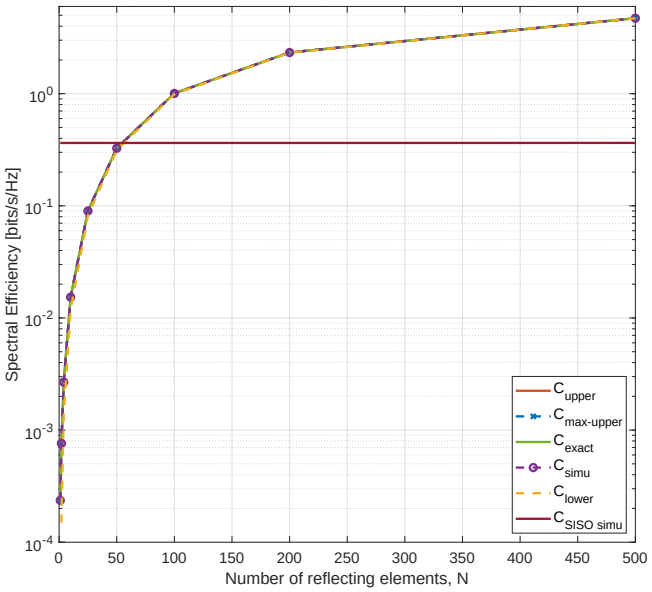
Regarding the distance between the source (i.e., the BS) and the LIS, we compare the schemes' spectral efficiencies with $N = 25, 50, 100, 250, 500$. It is worth mentioning that the user position is fixed, only the BS position changes. Figure 5.7 shows the results obtained for $b = 8$. We can see that the performance deteriorates as the distance increases. This phenomenon



(a)



(b)



(c)

Figure 5.5 – Spectral efficiency as a function of N for (a) $b = 1$ (b) $b = 4$ (c) $b = 10$.

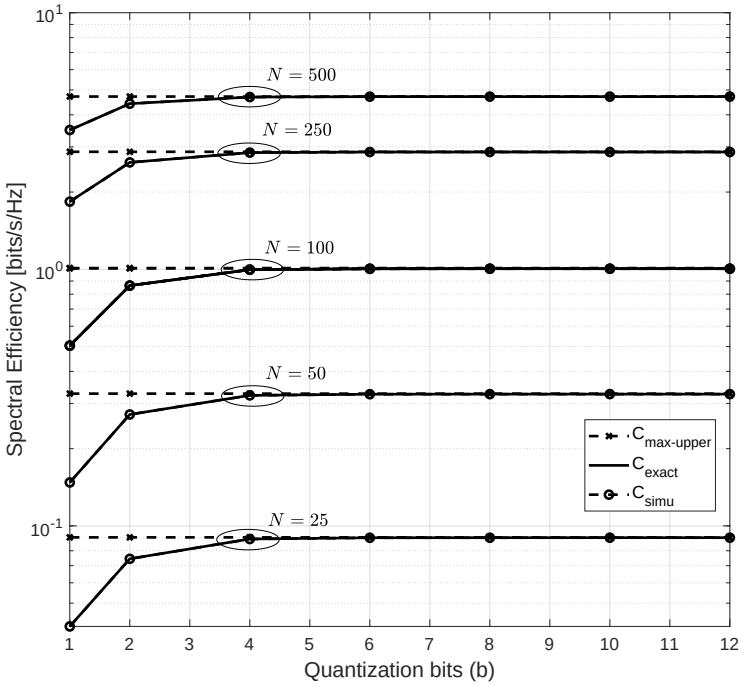


Figure 5.6 – Spectral efficiency as a function of b for different values of N.

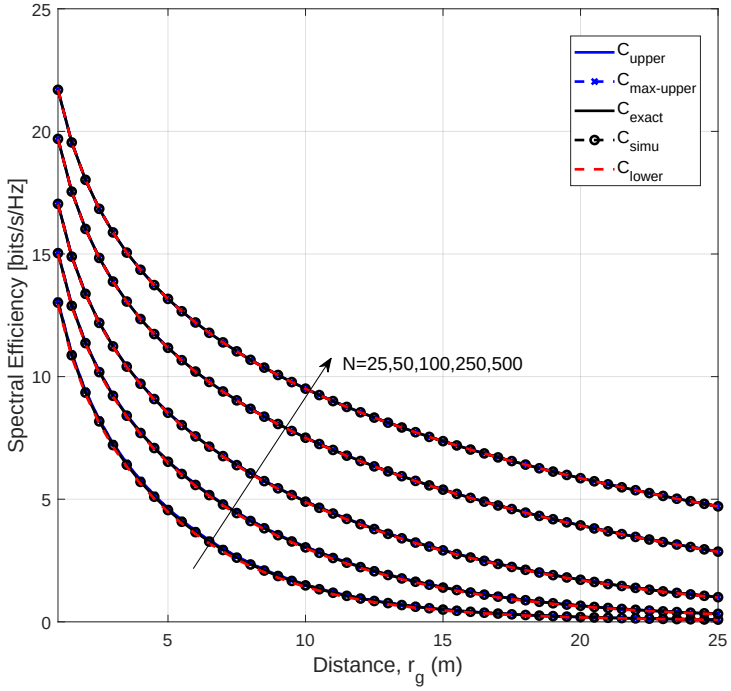


Figure 5.7 – Spectral efficiency as a function of the distance between source and LIS, considering $b = 8$.

was expected since the LIS is composed of only passive elements and there is no direct path between the source and the user. However, as already mentioned, it improves when the number of elements on LIS increases.

In its turn, Figure 5.8 shows how the spectral efficiency behaves as $\rho = P/N^{\alpha}$ varies for $\alpha = 3/2, 2, 5/2$. We consider $P = 100$ [dB] and $b \in \{1, 2, 4, 8, 10\}$. As expected and stated in Remark 14, for $\alpha = 2$ and as N increases, the capacity becomes constant no matter the

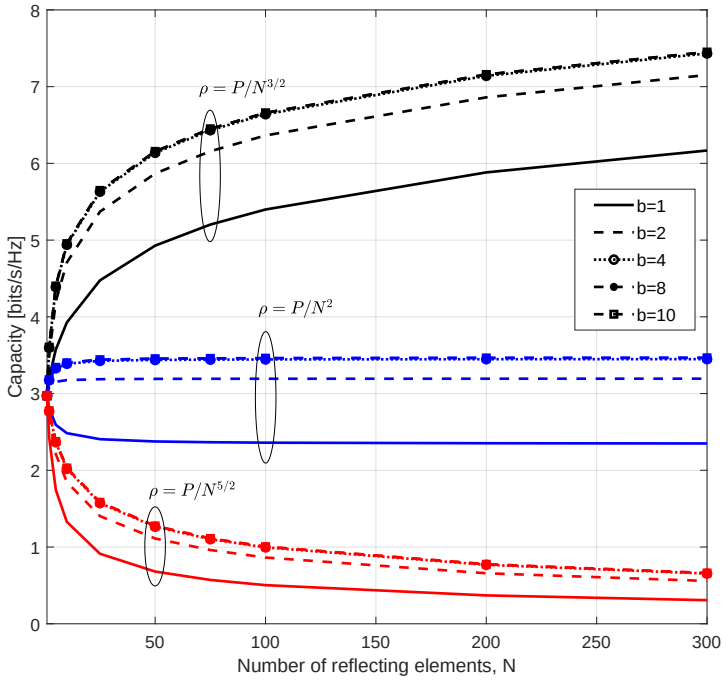


Figure 5.8 – Power scaling law for different α values.

number of reflecting elements. However, when $\alpha = 3/2$, the capacity grows logarithmically fast with N when $N \to \infty$ and tends to 0 when $\alpha = 5/2$ and $N \to \infty$. These results confirm that the transmit power can be reduced proportionally to N. We can also see that, although the capacity increases with the number of quantization bits, b, the performances for $b = 4$, $b = 8$, and $b = 10$ are very close.

In Figure 5.9, we show the required transmit power by the source needed to achieve fixed capacities of 1 and 2 bits/s/Hz, respectively, considering $\alpha = 2$. As expected and predicted by Remark 14, the transmit power can be reduced by approximately 6 [dB] by doubling the number of reflecting elements for both fixed capacities. We can also confirm that, in general, the LIS-assisted system outperforms the SISO system without LIS when N is approximately greater than 80, regardless of the number of quantization bits.

Figure 5.10 compares the average SNR as a function of the transmitted power obtained from the simulations, (5.11), and the SISO system without LIS. We consider $\rho = 50$ dB, $N \in \{25, 100, 200\}$ and $b \in \{1, 4, 8\}$. As can be confirmed, the relationship between the two parameters is linear, i.e., between the average SNR and the transmitted power. Additionally, we can notice that, for a given ρ, the average SNR improves as N increases. It is also worth mentioning that the LIS-assisted system outperforms the conventional one the higher the number of reflecting elements N is. Moreover, the influence of the number of quantization bits is insignificant, as long as $b > 1$.

Figures 5.11 and 5.12 shows the symbol error rate behavior for BPSK and QPSK, and 16-QAM, and 64-QAM modulation schemes considering $N = 25$, respectively. As expected, the modulations present a decreasing level of robustness as the number of symbols increases.

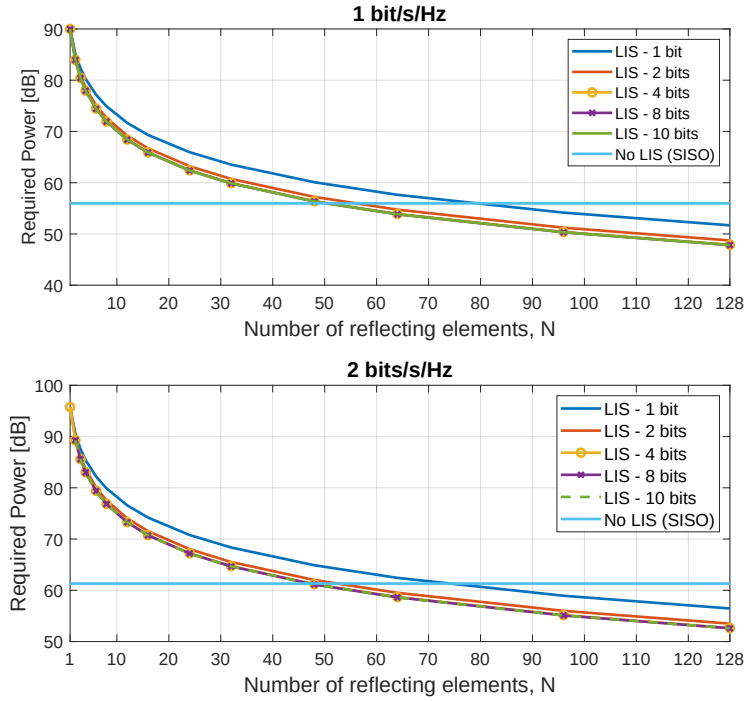


Figure 5.9 – Required power for $C = 1$ bit/s/Hz and $C = 2$ bits/s/Hz.

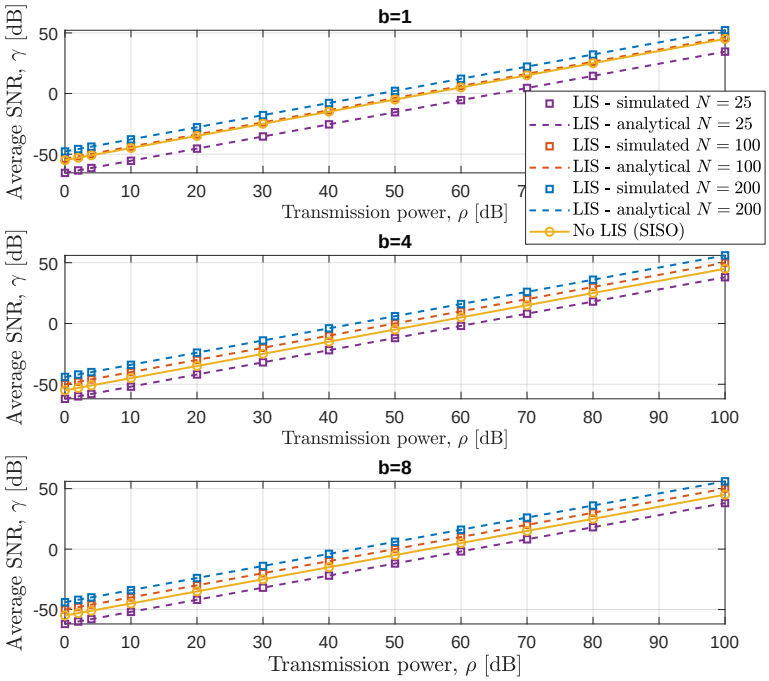


Figure 5.10 – Average SNR in function of transmission power for $b = 1$, $b = 4$ and $b = 8$.

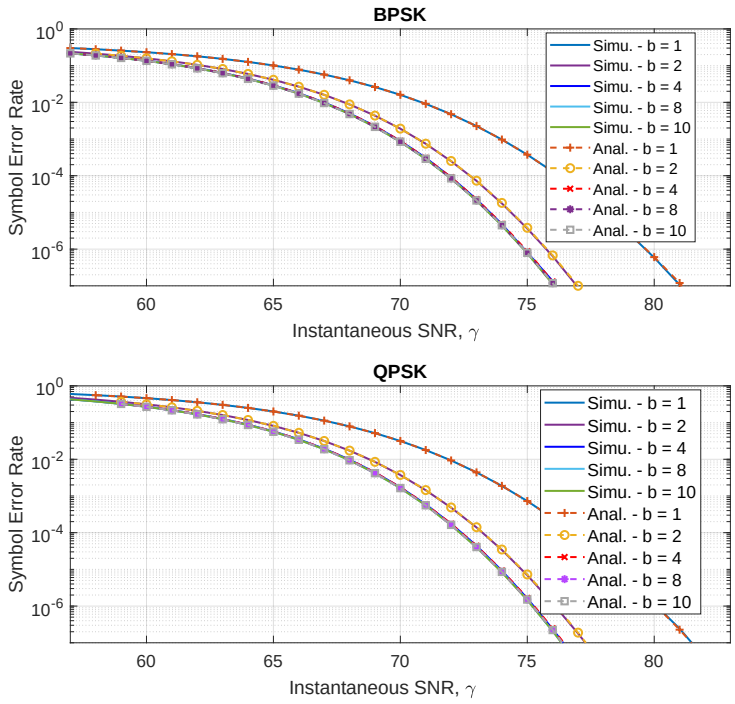


Figure 5.11 – Symbol Error Rate for BPSK and QPSK modulations for $N = 25$.

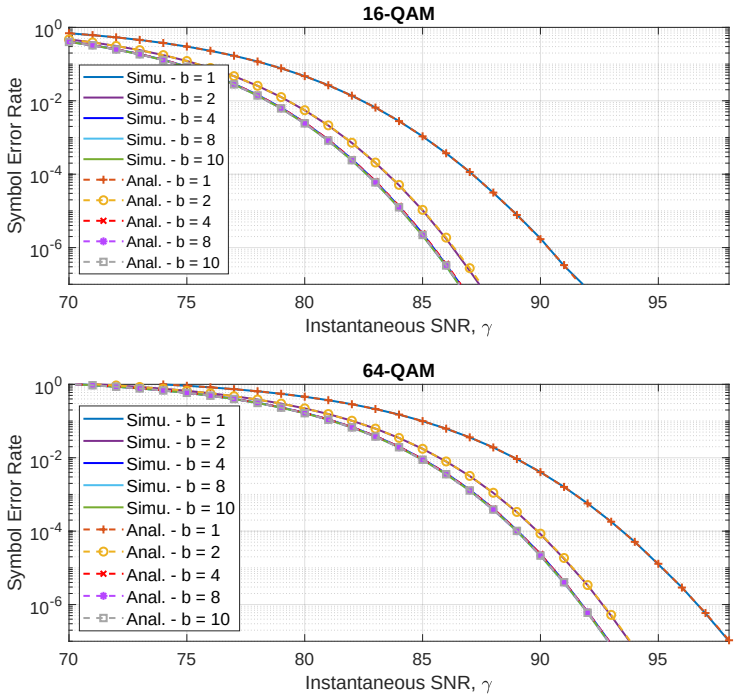


Figure 5.12 – Symbol Error Rate for 16-QAM and 64-QAM modulations for $N = 25$.

The most important thing to note here is the gap between the curves for 1, 2, and 4 bits. It gets to be almost 5 dB when the SNR is high. Although this gap exists, it is less pronounced when more bits are dedicated to phase quantization, and $b = 4$ is enough to guarantee a good performance.

For computational simplicity, in Figure 5.13, we present the simulated and analytical outage probabilities only for $N = 100$. We notice that the probability of achieving higher capacities

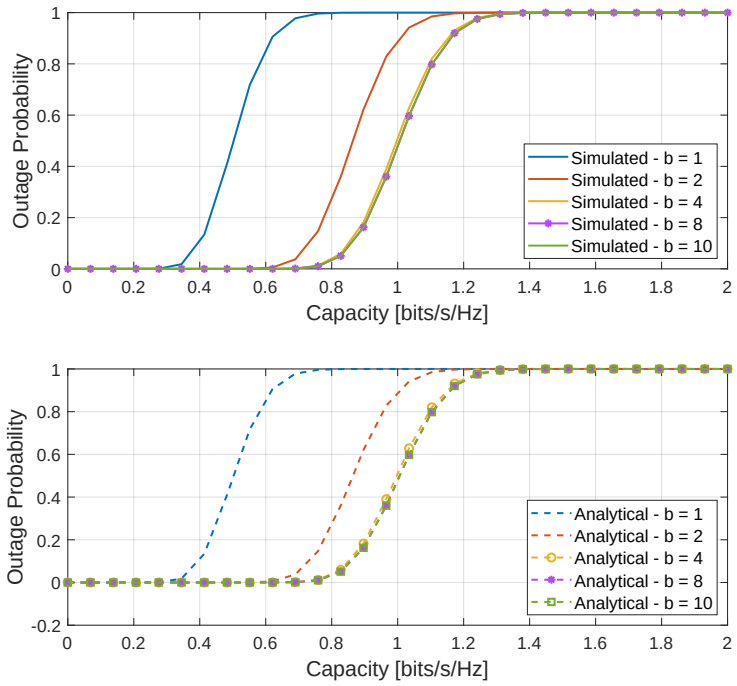


Figure 5.13 – Outage probability for $N = 100$ and $b \in \{1, 2, 4, 8, 10\}$.

increases with the number of quantization bits, b. Moreover, when considering different b values, it is possible to confirm our previous insight; $b = 4$ is enough for a good phase quantization.

Finally, Figure 5.14 shows the outage probability as we vary the number of quantization bits. For this simulation, we assume ρ equal to 80 [dB] and the capacity of 1 bit/s/Hz. By analyzing the results, it is possible to notice that the outage probability decreases as the number of quantization bits, b, increases and asymptotically approaches the lower bound given by an optimal LIS with perfect phase shifts, i.e., an infinite number of bits used to represent the phases. For all the three setups, $N = 2$, 4, and 8, the outage probability reaches its largest value when $b = 2$. When $N = 2$, the outage probability approaches the perfect phase bound as long as $b > 3$. We also notice that the gap between the curves gets larger as N increases. However, it reduces as the number of quantization bits, b, increases. Therefore, we conclude that the phase quantization errors do not significantly impact the outage probability performance as long as the number of bits is made large enough, which is an encouraging finding for the deployment of LIS-assisted systems.

5.5 Conclusion

In this chapter, we have done an in-depth analysis of a practical LIS-assisted Single-Input Single-Output (SISO) system. Since quantization errors are unavoidable, we evaluate the influence of bits number dedicated to the phase quantization on spectral efficiency, symbol error rate, and outage probability. We compare such a system performance with the conventional one

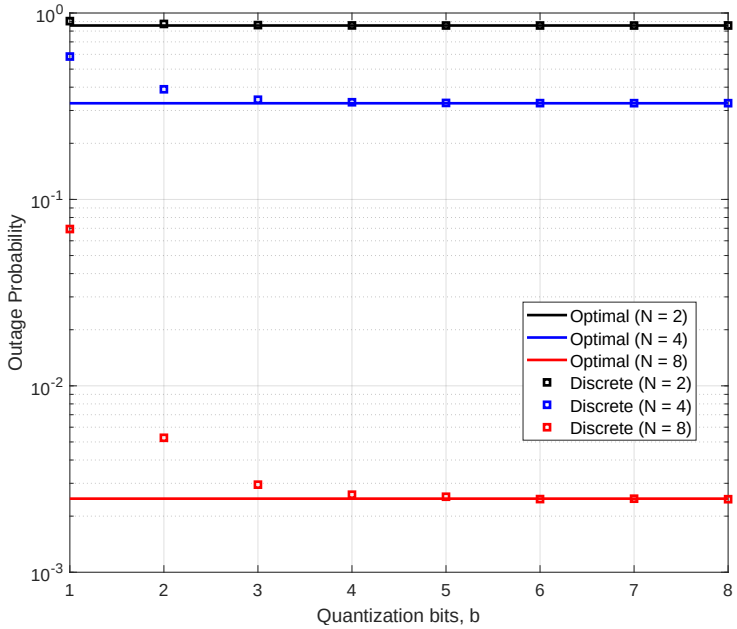


Figure 5.14 – Outage probability versus number of quantization bits, b, for different values of N.

without LIS through accurate closed-form expressions derived for each of these metrics. We have extended our analysis to power scaling law and the power required to achieve specific capacity. Not only is the influence of b verified, but also that of the number of LIS elements.

We can conclude that the performance improves as the number of LIS elements and bits increases. The LIS with approximately fifty elements and four dedicated bits for phase quantization outperforms the conventional system performance without LIS.

6

Concluding Remarks

This doctoral project has presented an overview of two key technologies of the mobile network over approximately five years. Our main concern was to analyze and model practical channels as close as possible to reality.

On one hand, we approach the performance of the massive MIMO of the fifth generation taking into account the interference inherent in it. We have derived an accurate approximation to SIR when the number of antennas at BS is or not finite. We consider path loss and lognormal shadowing.

As for the sixth generation, we got into topics related to intelligent surfaces. We initially compared the performance of such systems using Rician and Rayleigh channels. Due to the high computational cost, we evaluate the performance loss when BS does not have CSI and transmits the symbols with equal powers to all users.

With a more in-depth analysis, we derive analytical expressions for several parameters of practical LIS-assisted SISO systems as spectral efficiency, symbol error rate, and outage probability. We have assumed imperfections for the values of reflection coefficients considering even phase quantization errors.

We do not know how the future will be. But given the substantial improvements that massive MIMO and LIS technologies can offer, we can guarantee that the two are here to stay and coexist. So, as future work, we intend to:

- Generalize cases III and IV of the Chapter 3. We are aware that the models obtained are limited and only work well for the specified parameters;
- Deepen the knowledge related to LIS. It is worth remembering that the analyzes presented in Chapter 4 are preliminary since the analytical model to confirm our conclusions is missing. We need to derive equations for the general Rice case and then find the Rayleigh as a special case (when the Rice factor is zero);
- Encompass LIS in a neural network based on machine learning tools. In this way, the LIS would cease to be completely passive and would begin to feel and react alone to the environment with no sign of external control. It learns how to optimally interact with the incident signal given the channels at the active elements, extracting the current state information and transmitter/receiver locations [77];

- Insert the correlation matrix between the LIS elements in our analyzes would provide a more realistic analyses. Its construction is complex and takes into account the azimuth and elevation angles of each element. In practical LIS, the element spacing is smaller than $\lambda/2$ to approximate a continuously controllable surface, and i.i.d. fading cannot occur due to the planar nature [78];

- Improve the model by adding more antennas (at least) at the transmitter and also considering a direct link between transmitter and receiver. In this case, the channels have more dimensions, and the effect of phase noise becomes more important to study.

- Look further into the future. For example, millimeter waves will be increasingly explored, and as a consequence of a heterogeneous network, there will be no more borders in the cells. The so-called cell-free massive MIMO has many individually controllable antennas distributed over a wide area for simultaneously serving a small number of a user equipment.

Bibliography

[1] R. University, “Rice wins $2.4 million to study many-antenna wireless,” http://phys.org/news/2015-11-rice-million-many-antenna-wireless.html, nov., 2015.

[2] J. Y. D. S. J. Y. Z. Q. C. T. J. C. Wankai Tang, Xiang Li, “Wireless communications with programmable metasurface: Transceiver design and experimental results,” *China Communications*, vol. 16, no. 5, p. 46, 2019. [Online]. Available: http://www.cic-chinacommunications.cn/EN/abstract/article_895.shtml

[3] C. Huang, G. C. Alexandropoulos, A. Zappone, M. Debbah, and C. Yuen, “Energy efficient multi-user MISO communication using low resolution large intelligent surfaces,” *CoRR*, vol. abs/1809.05397, 2018.

[4] T. L. Marzetta, G. Caire, M. Debbah, I. Chih-Lin, and S. K. Mohammed, “Special issue on massive MIMO,” *Journal of Communications and Networks*, vol. 15, no. 4, pp. 333–337, Ago. 2013.

[5] J. C. M. Filho, “Desempenho e otimização de sistemas de comunicação sem-fio MIMO com elevado número de antenas,” Dissertação de Mestrado - Universidade Estadual de Londrina, 2014.

[6] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, “Five disruptive technology directions for 5G,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, Fev. 2014.

[7] T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[8] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan, and M. Zorzi, “Toward 6G networks: Use cases and technologies,” *IEEE Communications Magazine*, vol. 58, no. 3, pp. 55–61, 2020.

[9] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis, and T. L. Marzetta, “Massive MIMO is a reality-what is next?: Five promising research directions for antenna arrays,” *Digital Signal Processing*, vol. 94, pp. 3 – 20, 2019, special Issue on Source Localization in Massive MIMO. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1051200419300776

[10] A. Zappone, M. D. Renzo, and M. Debbah, "Wireless networks design in the era of deep learning: Model-based, ai-based, or both?" *IEEE Transactions on Communications*, vol. 67, pp. 7331–7376, 2019.

[11] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Network*, vol. 34, no. 3, pp. 134–142, 2020.

[12] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[13] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[14] "Designing 5G NR: The 3GPP release-15 global standard for a unified, more capable 5G air interface," https://www.qualcomm.com/media/documents/files/the-3gpp-release-15-5g-nr-design.pdf, accessed: 2018-09-05.

[15] A. Puglielli, N. Narevsky, P. Lu, T. Courtade, G. Wright, B. Nikolic, and E. Alon, "A scalable massive MIMO array architecture based on common modules," in *International Conference on Communication Workshop (ICCW)*, Jun. 2015.

[16] T. Bai and R. W. Heath, "Analyzing uplink SINR and rate in massive MIMO systems using stochastic geometry," *IEEE Transactions on Communications*, vol. 64, no. 11, pp. 4592–4606, Nov. 2016.

[17] J. C. M. Filho, C. Panazio, and T. Abrao, "Uplink performance of single-carrier receiver in massive MIMO with pilot contamination," *IEEE Access*, vol. 5, pp. 8669–8681, 2017.

[18] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th ed. McGraw-Hill Higher Education, 2002.

[19] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th ed. Boston: McGraw Hill, 2002. [Online]. Available: http://www.worldcat.org/search?qt=worldcat_org_all&q=0071226613

[20] J. C. S. S. Filho, P. Cardieri, and M. D. Yacoub, "Simple accurate lognormal approximation to lognormal sums," *Electronics Letters*, vol. 41, no. 18, pp. 1016–1017, Sep. 2005.

[21] P. Madhusudhanan, X. Li, Y. Liu, and T. X. Brown, "Stochastic geometric modeling and interference analysis for massive MIMO systems," in *International Symposium and Workshops on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, May 2013, pp. 15–22.

[22] C. Kong, C. Zhong, and Z. Zhang, "Performance of ZF precoder in downlink massive MIMO with non-uniform user distribution," *Journal of Communications and Networks*, vol. 18, no. 5, pp. 688–698, Oct. 2016.

[23] L. Atzori, A. Iera, and G. Morabito, "The internet of things: a survey," *Computer Networks, Elsevier*, vol. 54, no. 15, pp. 2787–2805, Oct. 2010.

[24] X. Tan, Z. Sun, J. M. Jornet, and D. Pados, "Increasing indoor spectrum sharing capacity using smart reflect-array," in *IEEE International Conference on Communications*, May 2016, pp. 1–6.

[25] K. David and H. Berndt, "6G vision and requirements: Is there any need for beyond 5G?" *IEEE Vehicular Technology Magazine*, vol. 13, no. 3, pp. 72–80, Sep. 2018.

[26] S. Hu, F. Rusek, and O. Edfors, "Beyond massive MIMO: The potential of positioning with large intelligent surfaces," *IEEE Transactions on Signal Processing*, vol. 66, no. 7, pp. 1761–1774, Apr. 2018.

[27] M. D. Renzo, M. Debbah, D. T. P. Huy, A. Zappone, M. Alouini, C. Yuen, V. Sciancalepore, G. C. Alexandropoulos, J. Hoydis, H. Gacanin, J. de Rosny, A. Bounceu, G. Lerosey, and M. Fink, "Smart radio environments empowered by AI reconfigurable meta-surfaces: An idea whose time has come," *CoRR*, vol. abs/1903.08925, 2019.

[28] E. Basar, M. D. Renzo, J. de Rosny, M. Debbah, M. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *CoRR*, vol. abs/1906.09490, 2019.

[29] S. V. Hum and J. Perruisseau-Carrier, "Reconfigurable reflectarrays and array lenses for dynamic antenna beam control: A review," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 1, pp. 183–198, Jan 2014.

[30] E. Bjornson, L. Sanguinetti, and M. Kountouris, "Deploying dense networks for maximal energy efficiency: Small cells meet massive MIMO," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 832–847, Apr. 2016.

[31] R. Mahapatra, Y. Nijsure, G. Kaddoum, N. Ul Hassan, and C. Yuen, "Energy efficiency tradeoff mechanism towards wireless green communication: A survey," *IEEE Communications Surveys Tutorials*, vol. 18, no. 1, pp. 686–705, Firstquarter 2016.

[32] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," 2018.

[33] D. R. Hunter and K. Lange, "A tutorial on mm algorithms," *The American Statistician*, vol. 58, no. 1, pp. 30–37, Dec. 2003.

[34] T. Hou, Y. Liu, Z. Song, X. Sun, Y. Chen, and L. Hanzo, "Reconfigurable intelligent surface aided noma networks," *ArXiv*, vol. abs/1912.10044, 2019.

[35] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 4157–4170, 2019.

[36] S. K. Jayaweera and H. V. Poor, "On the capacity of multiple-antenna systems in rician fading," *IEEE Transactions on Wireless Communications*, vol. 4, no. 3, pp. 1102–1111, 2005.

[37] Wolfram Research, "Mathematica," 2018. [Online]. Available: https://www.wolfram.com/mathematica/

[38] D. Chizhik, G. J. Foschini, M. J. Gans, and R. A. Valenzuela, "Keyholes, correlations and capacities of multi-element transmit and receive antennas," in *IEEE Vehicular Technology Conference*, vol. 1, May 2001, pp. 284–287 vol.1.

[39] S. K. Jayaweera and H. V. Poor, "On the capacity of multiple-antenna systems in Rician fading," *IEEE Transactions on Wireless Communications*, vol. 4, no. 3, pp. 1102–1111, May 2005.

[40] S. Hu, F. Rusek, and O. Edfors, "Beyond massive MIMO: The potential of data transmission with large intelligent surfaces," *IEEE Transactions on Signal Processing*, vol. 66, no. 10, pp. 2746–2758, May 2018.

[41] V. Arun and H. Balakrishnan, "Rfocus: Practical beamforming for small devices," 2019.

[42] "Greenerwave," http://greenerwave.com/, accessed: 2019-11-19.

[43] K. Ntontin, M. Di Renzo, J. Song, F. Lazarakis, J. de Rosny, D.-T. Phan-Huy, O. Simeone, R. Zhang, M. Debbah, G. Lerosey *et al.*, "Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison," *arXiv preprint arXiv:1908.08747*, 2019.

[44] N. C. Beaulieu, G. Farhadi, and Y. Chen, "A precise approximation for performance evaluation of amplify-and-forward multihop relaying systems," *IEEE Transactions on Wireless Communications*, vol. 10, no. 12, pp. 3985–3989, Dec. 2011.

[45] L. Sanguinetti, A. A. D'Amico, and Y. Rong, "A tutorial on the optimization of amplify-and-forward mimo relay systems," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 8, pp. 1331–1346, Sep. 2012.

[46] M. Jung, W. Saad, Y. Jang, G. Kong, and S. Choi, "Performance analysis of large intelligent surfaces (LISs): Asymptotic data rate and channel hardening effects," 2018.

[47] E. Basar, M. Di Renzo, J. de Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless Communications Through Reconfigurable Intelligent Surfaces," *arXiv e-prints*, p. arXiv:1906.09490, Jun 2019.

[48] M.-A. Badiu and J. P. Coon, "Communication Through a Large Reflecting Surface With Phase Errors," *arXiv e-prints*, p. arXiv:1906.10751, Jun 2019.

[49] S. Abeywickrama, R. Zhang, and C. Yuen, "Intelligent Reflecting Surface: Practical Phase Shift Model and Beamforming Optimization," *arXiv e-prints*, p. arXiv:1907.06002, Jul 2019.

[50] E. Basar, "Transmission through large intelligent surfaces: A new frontier in wireless communications," in *European Conference on Networks and Communications (EuCNC)*, Jun. 2019, pp. 112–117.

[51] J. Proakis and M. Salehi, *Digital Communications*. McGraw-Hill, 2008. [Online]. Available: https://books.google.com.br/books?id=ABSmAQAACAAJ

[52] W. Hou, T. Fujino, and T. Kojima, "Quantization error reduction for MIMO detection based on orthogonal lattices," in *IEICE Communications Express*, vol. 2, 2013, pp. 42–48.

[53] K. Kotera, O. Muta, and H. Furukawa, "Efficient nonlinear equalization scheme for MIMO constant envelope modulation receivers affected by quantization error," in *The International Conference on Information Network 2012*, 2012, pp. 275–279.

[54] Özgecan Özdogan, E. Björnson, and E. G. Larsson, "Intelligent reflecting surfaces: Physics, propagation, and pathloss modeling," 2019.

[55] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5394–5409, 2019.

[56] H. Han, J. Zhao, D. Niyato, M. D. Renzo, and Q. Pham, "Intelligent reflecting surface aided network: Power control for physical-layer broadcasting," in *IEEE International Conference on Communications (ICC)*, 2020, pp. 1–7.

[57] J. Ye, S. Guo, and M. Alouini, "Joint reflecting and precoding designs for ser minimization in reconfigurable intelligent surfaces assisted mimo systems," *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5561–5574, 2020.

[58] S. Abeywickrama, R. Zhang, and C. Yuen, "Intelligent reflecting surface: Practical phase shift model and beamforming optimization," in *IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.

[59] Y. Han, W. Tang, S. Jin, C. Wen, and X. Ma, "Large intelligent surface-assisted wireless communication exploiting statistical CSI," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 8238–8242, 2019.

[60] R. C. Ferreira, M. S. P. Facina, F. A. P. De Figueiredo, G. Fraidenraich, and E. R. De Lima, "Bit error probability for large intelligent surfaces under double-nakagami fading channels," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 750–759, 2020.

[61] D. Gesbert, H. Bolcskei, D. A. Gore, and A. J. Paulraj, "Outdoor MIMO wireless channels: models and performance prediction," *IEEE Transactions on Communications*, vol. 50, no. 12, pp. 1926–1934, 2002.

[62] Y. Ai, M. Cheffena, A. Mathur, and H. Lei, "On physical layer security of double Rayleigh fading channels for vehicular communications," *IEEE Wireless Communications Letters*, vol. 7, no. 6, pp. 1038–1041, 2018.

[63] J. Salo, H. M. El-Sallabi, and P. Vainikainen, "Impact of double-Rayleigh fading on system performance," in *International Symposium on Wireless Pervasive Computing*, 2006, p. 5.

[64] H. Shin and M. Z. Win, "MIMO diversity in the presence of double scattering," *IEEE Transactions on Information Theory*, vol. 54, no. 7, pp. 2976–2996, 2008.

[65] T. Wang, G. Chen, J. P. Coon, and M.-A. Badiu, "Study of intelligent reflective surface assisted communications with one-bit phase adjustments," 2020.

[66] T. Marzetta, E. Larsson, and H. Yang, *The fundamentals of Massive MIMO*. Cambridge University Press, 2016.

[67] M. D. Renzo, F. H. Danufane, X. Xi, J. de Rosny, and S. Tretyakov, "Analytical modeling of the path-loss for reconfigurable intelligent surfaces – anomalous mirror or scatterer?" 2020.

[68] Q. Wu and R. Zhang, "Beamforming optimization for intelligent reflecting surface with discrete phase shifts," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2019, pp. 7830–7833.

[69] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116 753–116 773, 2019.

[70] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, 7th ed. Elsevier/Academic Press, Amsterdam, 2007.

[71] "*N*IST Digital Library of Mathematical Functions," http://dlmf.nist.gov/, Release 1.0.25 of 2019-12-15, f. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F.

Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, eds. [Online]. Available: http://dlmf.nist.gov/

[72] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, ninth dover printing, tenth GPO printing ed. New York: Dover, 1964.

[73] Zhengdao Wang and G. B. Giannakis, "A simple and general parameterization quantifying performance in fading channels," *IEEE Transactions on Communications*, vol. 51, no. 8, pp. 1389–1398, 2003.

[74] F. A. Pereira de Figueiredo, C. F. Dias, E. Rodrigues de Lima, and G. Fraidenraich, "Capacity bounds for dense massive MIMO in a line-of-sight propagation environment," Dec. 2019. [Online]. Available: https://www.techrxiv.org/articles/preprint/Capacity_Bounds_for_Dense_Massive_MIMO_in_a_Line-of-Sight_Propagation_Environment/11395698/1

[75] F. A. P. De Figueiredo, F. A. C. M. Cardoso, I. Moerman, and G. Fraidenraich, "On the application of massive MIMO systems to machine type communications," *IEEE Access*, vol. 7, pp. 2589–2611, 2019.

[76] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Statist.*, vol. 22, no. 1, pp. 79–86, Mar. 1951. [Online]. Available: https://doi.org/10.1214/aoms/1177729694

[77] A. M. Elbir, A. Papazafeiropoulos, P. Kourtessis, and S. Chatzinotas, "Deep channel learning for large intelligent surfaces aided mm-wave massive mimo systems," *IEEE Wireless Communications Letters*, vol. 9, no. 9, pp. 1447–1451, 2020.

[78] E. Bjornson, "The end of independent Rayleigh fading," Oct. 2020. [Online]. Available: http://ma-mimo.ellintech.se/2020/10/15/the-end-of-independent-rayleigh-fading/

[79] Q. Zhang, S. Jin, K. Wong, H. Zhu, and M. Matthaiou, "Power scaling of uplink massive MIMO systems with arbitrary-rank channel means," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 966–981, 2014.

[80] Wolfram Research, "The wolfram functions site," Oct. 2001. [Online]. Available: https://functions.wolfram.com/HypergeometricFunctions/Hypergeometric1F1/03/01/01/0001/

[81] J. Shynk, *Probability, Random Variables, and Random Processes: Theory and Signal Processing Applications*. Wiley, 2012. [Online]. Available: https://books.google.com.br/books?id=lJsTPzo37tAC

[82] W. R. Inc., "Mathematica, Version 12.1," champaign, IL, 2020. [Online]. Available: https://www.wolfram.com/mathematica

A.1 Mean and Variance of C

The mean value of C is given by

$$\mu_C = E[C] = E\left[\sum_{k=1}^{n} \frac{1}{n}|H_{12k}|\cos\theta_k\right]. \tag{A.1}$$

Assuming that all coefficients are independent and have the same statistics, the following can be written

$$\begin{aligned} \mu_C &= n \times \frac{1}{n} \times E\left[|H_{12k}|\cos\theta_k\right] \\ &= \mu_1\mu_2\left(\frac{I_1(\kappa)}{I_0(\kappa)}\right), \end{aligned} \tag{A.2}$$

in which $\mu_1 = E[H_{1k}]$ and $\mu_2 = E[H_{2k}]$. The variance of C is given by

$$\sigma_C^2 = \text{var}\left(\sum_{k=1}^{n} \frac{1}{n}|H_{12k}|\cos\theta_k\right). \tag{A.3}$$

Note that the variance computation in (A.3) involves the product of two independent variables. Let the variables u and v be independent. Therefore, the variance of the product is given by $\text{var}(uv) = \text{var}(u)\text{var}(v) + \text{var}(u)E[v]^2 + \text{var}(v)E[u]^2$ [19]. Additionally, since all the variables are independent, the following can be written

$$\begin{aligned} \sigma_C^2 &= \frac{1}{n}\text{var}\left(|H_{12k}|\cos\theta_k\right) \\ &= \frac{1}{n}\left(\sigma_{12}^2\sigma_{C_k}^2 + \sigma_{12}^2\mu_{C_k}^2 + \sigma_{C_k}^2\mu_{12}^2\right), \end{aligned} \tag{A.4}$$

whose parameters can be calculated as follows

$$\sigma_{12}^2 = \sigma_1^2\sigma_2^2 + \mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2, \tag{A.5}$$

in which $\sigma_1^2 = \text{var}(H_{1k})$ and $\sigma_2^2 = \text{var}(H_{2k})$,

$$\sigma_{C_k}^2 = E\left[cos^2\theta_k\right] - E^2\left[\cos\theta_k\right], \tag{A.6}$$

$$\mu_{C_k} = E\left[\cos\theta_k\right] = \alpha_1 = \frac{I_1(\kappa)}{I_0(\kappa)}, \tag{A.7}$$

$$E\left[cos^2\theta_k\right] = \frac{1}{2} + \frac{1}{2}E\left[\cos 2\theta_k\right] = \frac{1}{2} + \frac{1}{2}\alpha_2. \tag{A.8}$$

and

$$\sigma^2_{C_k} = \frac{1}{2} + \frac{1}{2}\frac{I_2(\kappa)}{I_0(\kappa)} + \left(\frac{I_1(\kappa)}{I_0(\kappa)}\right)^2. \tag{A.9}$$

Therefore, the variance of C is defined as

$$\begin{aligned}\sigma^2_C = \frac{1}{2n}\sigma^2_{12}\left(1 + \frac{I_2(\kappa)}{I_0(\kappa)} + 4\left[\frac{I_1(\kappa)}{I_0(\kappa)}\right]^2\right) + \dots \\ \cdots + \mu_1\mu_2\frac{1}{2n}\left(1 + \frac{I_2(\kappa)}{I_0(\kappa)}\right).\end{aligned} \tag{A.10}$$

A.2 Mean and Variance of S

The mean value of S can be written as

$$\begin{aligned}\mu_S &= E[S] = E\left[\sum_{k=1}^{n}\frac{1}{n}|H_{12k}|\sin\theta_k\right] \\ &= \frac{1}{n}\times n\times\mu_1\mu_2 E[\sin\theta_k],\end{aligned} \tag{A.11}$$

Using the same assumptions as in the previous case, the mean value of S can be calculated as

$$\mu_{S_k} = E\left[\sin\theta_k\right] = \beta_1 = 0. \tag{A.12}$$

Therefore,

$$\mu_S = 0 \tag{A.13}$$

On the other hand, the variance of S can be computed as

$$\begin{aligned}\sigma^2_S &= \mathrm{var}\left(\sum_{k=1}^{n}\frac{1}{n}|H_{12k}|\sin\theta_k\right) \\ &= \frac{1}{n}\left(\sigma^2_{12}\sigma^2_{S_k} + \sigma^2_{12}\mu^2_{S_k} + \sigma^2_{S_k}\mu^2_{12}\right),\end{aligned} \tag{A.14}$$

whose parameters can be calculated as follows

$$\begin{aligned}\sigma^2_{S_k} &= E\left[sin^2\theta_k\right] - E^2\left[\sin\theta_k\right] \\ &= \frac{1}{2} - \frac{1}{2}\frac{I_2(\kappa)}{I_0(\kappa)}\end{aligned} \tag{A.15}$$

$$E\left[sin^2\theta_k\right] = \frac{1}{2} - \frac{1}{2}E\left[\cos 2\theta_k\right] = \frac{1}{2} - \frac{1}{2}\alpha_2.$$

Therefore, the variance of S is defined as

$$\sigma^2_S = \frac{1}{2n}\sigma^2_{12}\left(1 - \frac{I_2(\kappa)}{I_0(\kappa)}\right) + \frac{1}{2n}(\mu_1\mu_2)^2\left(1 - \frac{I_2(\kappa)}{I_0(\kappa)}\right). \tag{A.16}$$

A.2.1 Mean of the product of C and S

This appendix shows that C and S are uncorrelated. Considering a possible correlation between C and S, the bivariate Gaussian joint distribution can be written as

$$f_{C,S}(x,y) = \frac{1}{2\pi\sigma_C\sigma_S\sqrt{1-\rho^2}} e^{-\frac{1}{2\sqrt{1-\rho^2}}\left(\frac{(x-\mu_C)^2}{\sigma_C^2}+\frac{(y-\mu_S)^2}{\sigma_S^2}-\frac{2\rho(x-\mu_C)(y-\mu_S)}{\sigma_C\sigma_S}\right)} \tag{A.17}$$

where ρ is defined as [19]

$$\rho = \frac{E\left[CS\right] - E\left[C\right]E\left[S\right]}{\sigma_C\sigma_S} \tag{A.18}$$

As it has been calculated from (A.13), the term $E[S] = 0$, that is, the second term of the numerator of (A.18) is zero. Therefore in order to prove that $\rho = 0$, we need to compute the mean of the product between C and S and show that is also zero.

Departing from (4.19), the mean of the product between C and S can be written as

$$E\left[CS\right] = E\left[\sum_{i=1}^{n}\sum_{j=1}^{n} R_iR_j\cos\left(\theta_i\right)\sin\left(\theta_j\right)\right] \tag{A.19}$$

Since R_k is independent of θ_k for all k, then

$$\begin{aligned} E[CS] = 2\sum_{i=1}^{n}\sum_{j>i}^{n} E\left[R_iR_j\right]E\left[\cos\left(\theta_i\right)\sin\left(\theta_j\right)\right] + \\ \sum_{k=1}^{n} E\left[R_i^2\right]E\left[\sin\left(\theta_i\right)\cos\left(\theta_i\right)\right] \end{aligned} \tag{A.20}$$

In the sequel, the following equalities are proven

$$E\left[\cos\left(\theta_i\right)\sin\left(\theta_j\right)\right] = 0$$

and

$$E\left[\sin\left(\theta_i\right)\cos\left(\theta_i\right)\right] = 0$$

Therefore (A.20) will be null. Using the very definition of the mean, the term $E\left[\cos\left(\theta_i\right)\sin\left(\theta_j\right)\right]$ can be computed as

$$E[\cos(\theta_i)\sin(\theta_j)] = \int_0^{2\pi}\int_0^{2\pi}\cos(\theta_i)\sin(\theta_j)\frac{e^{\kappa\cos(\theta_i)}}{2\pi I_0(\kappa)}\frac{e^{\kappa\cos(\theta_j)}}{2\pi I_0(\kappa)}d\theta_i d\theta_j = 0, \tag{A.21}$$

in the same way, the mean with respect to θ_i in the second term of (A.20), can be calculated as

$$E[\cos(\theta_i)\sin(\theta_i)] = \int_0^{2\pi}\frac{\cos(\theta_i)\sin(\theta_i)}{2\pi I_0(\kappa)}e^{\kappa\cos(\theta_i)}d\theta_i = 0, \tag{A.22}$$

therefore $E[CS] = 0$ and consequently ρ, given in (A.18)), will be $\rho = 0$. From this, (A.17) can be written as in (**??**).

A.2.2 Variance of C and S for Case IV

$$\sigma_C^2 = \frac{1}{768n^2\sigma_R^4}\sum_{i=1}^{n}\Bigg(192\left((\beta_{\min}-1)^2 4^{-k}\,{}_2F_1\left(\frac{1}{2}-k,-k;2;1\right)+2\beta_{\min}\Big(\beta_{\min}-(\beta_{\min}-1)2^{-k}\right.$$
$${}_2F_1\left(\frac{1-k}{2},-\frac{k}{2};2;1\right)+\frac{4^{-k}(\beta_{\min}-1)}{\pi(2k+1)^2}\Big(768(\beta_{\min}-1)(2k+1)\cos^2(\alpha_i+\delta_i)$$
$${}_3F_2\left(-\frac{1}{2},\frac{1}{2},1;k+1,k+\frac{3}{2};1\right)+512(\beta_{\min}-1)k(2k+1)^2\cos^2(\alpha_i+\delta_i)\,{}_3F_2\left(1,\frac{1}{2}-k,1-k;\frac{3}{2},\frac{5}{2};1\right)+$$
$$\frac{48(\beta_{\min}-1)(2k+1)\sin^2(\alpha_i+\delta_i)\left(\frac{16(k+1)\,{}_3F_2\left(\frac{1}{2},\frac{3}{2},2;k+2,k+\frac{5}{2};1\right)}{2k+3}-\frac{(2\pi k+\pi)\,{}_3F_2\left(\frac{1}{2},1,\frac{3}{2};\frac{k}{2}+\frac{3}{2},\frac{k}{2}+2;1\right)^2}{(k+2)^2}\right)}{(k+1)^2}+$$
$$\frac{1}{(k+1)(k+2)}24\pi(\beta_{\min}-1)(2k+1)\,{}_3F_2\left(\frac{1}{2},1,\frac{3}{2};\frac{k}{2}+\frac{3}{2},\frac{k}{2}+2;1\right)\sin^2(\alpha_i+\delta_i)$$
$$\left((8k+4)\,{}_3F_2\left(1,\frac{1}{2}-\frac{k}{2},-\frac{k}{2};\frac{1}{2},\frac{3}{2};1\right)+\pi k(k+2)\,{}_2F_1\left(\frac{1-k}{2},-\frac{k}{2};2;1\right)\right)-\frac{1}{2k+3}$$
$$2^{k+7}k(4k(k+2)+3)\beta_{\min}\left(4(2k+1)\cos^2(\alpha_i+\delta_i)\,{}_3F_2\left(1,\frac{1}{2}-\frac{k}{2},1-\frac{k}{2};\frac{3}{2},\frac{5}{2};1\right)+\sin^2(\alpha_i+\delta_i)\right.$$
$$\left(8(2k+1)\,{}_3F_2\left(2,\frac{1}{2}-\frac{k}{2},1-\frac{k}{2};\frac{3}{2},\frac{5}{2};1\right)+3\pi(k-1)\,{}_2F_1\left(\frac{1-k}{2},-\frac{k}{2};2;1\right)\right)+\frac{1}{k+1}$$
$$\left(-3\beta_{\min}2^{k+9}(2k+1)^2\cos^2(\alpha_i+\delta_i)\,{}_3F_2\left(-\frac{1}{2},\frac{1}{2},1;\frac{k}{2}+1,\frac{k}{2}+\frac{3}{2};1\right)+(3\pi(\beta_{\min}-1)k\,(64(2k-1)\right.$$
$$(2k+1)(4k-1)\,{}_2F_1\left(1-k,\frac{3}{2}-k;2;1\right)-\pi^2k(k+1)(k+2)\,{}_2^2F_1\left(\frac{1-k}{2},-\frac{k}{2};2;1\right)^2+8(2k+1)$$
$$\left(-3\pi^2(\beta_{\min}-1)k(k+2)(k+1)\,{}_2F_1\left(\frac{1-k}{2},-\frac{k}{2};2;1\right)\,{}_3F_2\left(1,\frac{1}{2}-\frac{k}{2},-\frac{k}{2};\frac{1}{2},\frac{3}{2};1\right)-6(k+1)\right.$$
$$\pi(\beta_{\min}-1)(2k+1)\,{}_3F_2\left(1,\frac{1}{2}-\frac{k}{2},-\frac{k}{2};\frac{1}{2},\frac{3}{2};1\right)^2+16(2k+1)\quad 8(\beta_{\min}-1)k(k+1)$$
$${}_3F_2\left(2,\frac{1}{2}-k,1-k;\frac{3}{2},\frac{5}{2};1\right)-3\sqrt{\pi}\beta_{\min}\Gamma(k+2)\,{}_3\tilde{F}_2\left(\frac{1}{2},\frac{3}{2},2;\frac{k+4}{2},\frac{k+5}{2};1\right)\sin^2(\alpha_i+\delta_i)$$
(A.23)

$$\sigma_S^2 = \frac{1}{768n^2\sigma_R^4}\sum_{i=1}^{n}\left(4^{-k}3\left(\beta_{\min}^2 2^{2k+7} + \frac{64(\beta_{\min}-1)^2\,{}_2F_1\left(\frac{1}{2}-k,-k;2;1\right)\left(2k\cos^2(\alpha_i+\delta_i)+4k+1\right)}{4k+1}\right.\right.$$
$$(2k-1)+\frac{(\beta_{\min}-1)\,{}_2F_1\left(\frac{1-k}{2},-\frac{k}{2};2;1\right)}{(2k+1)^2}\left(\beta_{\min}\left(-2^{k+7}\right)(2k+1)\left((k-1)k\cos^2(\alpha+\delta)+2k+1\right)\right.\pi^2$$
$$(\beta_{\min}-1)\left(-k^2\right)(k+2)^2\,{}_2F_1\left(\frac{1-k}{2},-\frac{k}{2};2;1\right)\cos^2(\alpha i+\delta i) - 3(\beta_{\min}-1)^2 4^{2-k}\cos^2(\alpha_i+\delta_i)$$
$${}_3F_2\left(1,\frac{1}{2}-\frac{k}{2},-\frac{k}{2};\frac{1}{2},\frac{3}{2};1\right)^2 + 3\sqrt{\pi}(\beta_{\min}-1)^2 8^{1-k}\cos^2(\alpha_i+\delta_i)\,{}_3F_2\left(1,\frac{1}{2}-\frac{k}{2},-\frac{k}{2};\frac{1}{2},\frac{3}{2};1\right)$$
$$\Gamma(k+1)\,{}_3\tilde{F}_2\left(\frac{1}{2},1,\frac{3}{2};\frac{k+3}{2},\frac{k+4}{2};1\right) - \frac{\sqrt{\pi}2^k k(k+2)\,{}_2F_1\left(\frac{1-k}{2},-\frac{k}{2};2;1\right)}{2k+1} + \frac{1}{\pi}2^{-4k}(\beta_{\min}-1)$$
$$\left((\beta_{\min}-1)4^{k+5}k\cos^2(\alpha_i+\delta_i){}_3F_2\left(2,\frac{1}{2}-k,1-k;\frac{3}{2},\frac{5}{2};1\right) - \beta_{\min}8^{k+3}k(\cos(2(\alpha_i+\delta_i))+1)\right.$$
$${}_3F_2\left(2,\frac{1}{2}-\frac{k}{2},1-\frac{k}{2};\frac{3}{2},\frac{5}{2};1\right) - 3\pi^2(\beta_{\min}-1)\Gamma(k+1)^2\cos^2(\alpha_i+\delta_i)\,{}_3\tilde{F}_2\left(\frac{1}{2},1,\frac{3}{2};\frac{k+3}{2},\frac{k+4}{2};1\right)^2 +$$
$$(\beta_{\min}-1)2^{2k+9}k\sin^2(\alpha_i+\delta_i){}_3F_2\left(1,\frac{1}{2}-k,1-k;\frac{3}{2},\frac{5}{2};1\right) - \beta_{\min}8^{k+3}k\sin^2(\alpha_i+\delta_i)$$
$${}_3F_2\left(1,\frac{1}{2}-\frac{k}{2},1-\frac{k}{2};\frac{3}{2},\frac{5}{2};1\right) + 192\sqrt{\pi}(\beta_{\min}-1)\Gamma(2k+1)\left(\cos^2(\alpha_i+\delta_i)\,{}_3\tilde{F}_2\left(\frac{1}{2},\frac{3}{2},2;k+2,k+\frac{5}{2};1\right)+\right.$$
$$\left.2\sin^2(\alpha_i+\delta_i)\,{}_3\tilde{F}_2\left(-\frac{1}{2},\frac{1}{2},1;k+1,k+\frac{3}{2};1\right)\right) + 2^{k+1}3\sqrt{\pi}\Gamma(k+1)\left(\frac{\pi^2(\beta_{\min}-1)k(k+2)\,\cos^2(\alpha_i+\delta_i)}{2k+1}\right)$$
$${}_2F_1\left(\frac{1-k}{2},-\frac{k}{2};2;1\right)_3\tilde{F}_2\left(\frac{1}{2},1,\frac{3}{2};\frac{k+3}{2},\frac{k+4}{2};1\right)2^{k+6}\beta_{\min}\left(\cos^2(\alpha_i+\delta_i)\right.$$
$${}_3\tilde{F}_2\left(\frac{1}{2},\frac{3}{2},2;\frac{k+4}{2},\frac{k+5}{2};1\right) + \left.\left.\left.\left.\left.2\sin^2(\alpha_i+\delta_i){}_3\tilde{F}_2\left(-\frac{1}{2},\frac{1}{2},1;\frac{k+2}{2},\frac{k+3}{2};1\right)\right)\right)\right)\right) \quad \text{(A.24)}$$

B.1 Parameters κ and θ

To find the parameters κ and θ for the approximated PDF of λ, we first need to define the following Lemmas.

Lemma 2

$$\left|\sum_{n=1}^{N} z_n e^{j\theta_n}\right|^2 = \sum_{n=1}^{N} z_n^2 + 2\sum_{m=1}^{N}\sum_{n=m+1}^{N} z_m z_n \cos(\theta_m - \theta_n). \tag{B.1}$$

Proof 1 *This identity is straightforwardly found by expanding the summation terms on its left side.*

Lemma 3 *If* $\boldsymbol{X} \sim \mathcal{CN}(\boldsymbol{0}_M, \sigma_X^2 \boldsymbol{I}_M)$*, then* $Y = |X|$ *is a Rayleigh random variable with PDF given by*

$$f_Y(y) = \frac{2y}{\sigma_X^2} e^{-\frac{y^2}{\sigma_X^2}}, y \geq 0. \tag{B.2}$$

Proof 2 *The proof for this Lemma is given in [19].*

Lemma 4 *If* Y *is a Rayleigh random variable with PDF defined by* (B.2)*, then, its 4 first moments are given by*

$$\mathbb{E}[Y] = \int_0^\infty y f_Y(y) dy = \frac{\sigma_X\sqrt{\pi}}{2}, \tag{B.3}$$

$$\mathbb{E}\left[Y^2\right] = \int_0^\infty y^2 f_Y(y) dy = \sigma_X^2, \tag{B.4}$$

$$\mathbb{E}\left[Y^3\right] = \int_0^\infty y^3 f_Y(y) dy = \frac{3\sigma_X^3\sqrt{\pi}}{4}, \tag{B.5}$$

$$\mathbb{E}\left[Y^4\right] = \int_0^\infty y^4 f_Y(y) dy = 2\sigma_X^4. \tag{B.6}$$

Lemma 5 *If X is a uniform random variable with PDF given by*

$$f_X(x) = \begin{cases} \frac{a}{2\pi}, & -\frac{\pi}{a} \leq x \leq \frac{\pi}{a}, \\ 0, & otherwise, \end{cases} \tag{B.7}$$

then $Y = -X$ has the same PDF as X, which was defined in (B.7).

Proof 3 *This can be straightforwardly proved by noticing that the PDF of X is symmetrical around 0.*

Lemma 6 *If θ_m and θ_n are independent and identically distributed uniform random variables with PDF given by* (B.7)*, then $Y = \theta_m + \theta_n$ has the following PDF*

$$f_Y(y) = \begin{cases} \frac{a}{2\pi}\left(1 + \frac{a}{2\pi}y\right), & -\frac{2\pi}{a} \leq y \leq 0, \\ \frac{a}{2\pi}\left(1 - \frac{a}{2\pi}y\right), & 0 < y \leq \frac{2\pi}{a}, \\ 0, & otherwise. \end{cases} \tag{B.8}$$

Proof 4 *From the theory, we know that the sum of two random variables equals the convolution of $f_{\theta_m}(\theta_m)$ and $f_{\theta_n}(\theta_n)$ is*

$$f_Y(y) = \int_{-\infty}^{\infty} f_{\theta_m}(y - \theta_n) f_{\theta_n}(\theta_n) d\theta_n. \tag{B.9}$$

Therefore, $f_Y(y)$ is defined as

$$f_Y(y) = \begin{cases} \int_{-\frac{\pi}{a}}^{\frac{\pi}{a}+y} \frac{a^2}{4\pi^2} d\theta_n, & -\frac{2\pi}{a} \leq y < 0, \\ \int_{-\frac{\pi}{a}+y}^{\frac{\pi}{a}} \frac{\beta_g}{4\pi^2} d\theta_n, & 0 \leq y \leq \frac{2\pi}{a}, \\ 0, & otherwise, \end{cases} \tag{B.10}$$

which concludes the proof.

Lemma 7 *If the PDF of the sum of two independent and identically distributed uniform random variables is given by* (B.8)*, then*

$$\mathbb{E}\left[\cos(\theta_m - \theta_n)\right] = \frac{a^2 \sin^2(\frac{\pi}{a})}{\pi^2}. \tag{B.11}$$

Proof 5 *By using Lemma 5, we can rewrite* (B.11) *as $\mathbb{E}\left[\cos(\theta_m + \theta_n)\right]$, then applying Lemma 6 we have*

$$\mathbb{E}\left[\cos(\theta_m + \theta_n)\right] = \mathbb{E}\left[\cos(y)\right] = \int_{-\frac{2\pi}{a}}^{0} \cos(y) \frac{a}{2\pi}\left(1 + \frac{a}{2\pi}y\right) dy + \int_{-\frac{2\pi}{a}}^{0} \cos(y) \frac{a}{2\pi}\left(1 - \frac{a}{2\pi}y\right) dy. \tag{B.12}$$

Solving the two integrals in (B.12) *concludes the proof.*

Lemma 8 *If the PDF of the sum of two independent and identically distributed uniform random variables is given by* (B.8)*, then*

$$\mathbb{E}\left[\cos^2(\theta_m - \theta_n)\right] = \frac{8\pi^2 + a^2 - a^2\cos^2(\frac{4\pi}{a})}{16\pi^2}. \tag{B.13}$$

Proof 6 *By using Lemma 5 we can rewrite* (B.11) *as* $\mathbb{E}\left[\cos^2(\theta_m + \theta_n)\right]$*, then applying Lemma 6 we have*

$$\mathbb{E}\left[\cos^2(\theta_m + \theta_n)\right] = \mathbb{E}\left[\cos^2(y)\right] = \int_{-\frac{2\pi}{a}}^{0} \cos^2(y)\frac{a}{2\pi}\left(1 + \frac{a}{2\pi}y\right)dy + \int_{-\frac{2\pi}{a}}^{0} \cos^2(y)\frac{a}{2\pi}\left(1 - \frac{a}{2\pi}y\right)dy. \tag{B.14}$$

Solving the two integrals in (B.14) *concludes the proof.*

Lemma 9 *If X is a uniform random variable with PDF given by* (B.7)*, then the PDF of $Y = 2X$ is given by*

$$f_Y(y) = \frac{a}{4\pi}, -\frac{2\pi}{a} \leq y \leq \frac{2\pi}{a}. \tag{B.15}$$

Proof 7 *This is proved by using the standard transformation of random variables.*

Lemma 10 *If θ_l, θ_m and θ_n are independent and identically distributed uniform random variables with PDF given by* (B.7)*, then $Y = 2\theta_l - (\theta_m + \theta_n)$ has the following PDF*

$$f_Y(y) = \begin{cases} \frac{a}{2\pi} + \frac{a^2 y}{4\pi^2} + \frac{a^3 y^2}{32\pi^3}, & -\frac{4\pi}{a} \leq y \leq -\frac{2\pi}{a}, \\ \frac{a}{4\pi} - \frac{a^3 y^2}{32\pi^3}, & -\frac{2\pi}{a} < y \leq 0, \\ \frac{a}{4\pi} - \frac{a^3 y^2}{32\pi^3}, & 0 < y \leq \frac{2\pi}{a}, \\ \frac{a}{2\pi} - \frac{a^2 y}{4\pi^2} + \frac{a^3 y^2}{32\pi^3}, & \frac{2\pi}{a} < y \leq \frac{4\pi}{a}, \\ 0, & \textit{otherwise.} \end{cases} \tag{B.16}$$

Proof 8 *We start by remembering that we know the PDF of $W = 2\theta_l$ and of $Z = \theta_m + \theta_n$, which are given by* (B.15) *and* (B.8)*, respectively. Next by applying Lemma 5, we can re-write Y as $Y = Z + W$, which is the sum of two independent random variables. Therefore, the PDF of Y is the convolution between the PDFs of W and Z, which is defined as*

$$f_Y(y) = \int_{-\infty}^{\infty} f_W(y - z) f_Z(z) dz. \tag{B.17}$$

Therefore, $f_Y(y)$ is defined as

$$f_Y(y) = \begin{cases} \int_{-\frac{2\pi}{a}}^{\frac{2\pi}{a}+y} \frac{a^2}{8\pi^2}\left(1+\frac{a}{2\pi}z\right)dz, & -\frac{4\pi}{a} \le y < -\frac{2\pi}{a}, \\ \int_{0}^{\frac{2\pi}{a}+y} \frac{a^2}{8\pi^2}\left(1-\frac{a}{2\pi}z\right)dz + \int_{-\frac{\pi}{a}+y}^{\frac{\pi}{a}} \frac{a^2}{8\pi^2}\left(1+\frac{a}{2\pi}z\right)dz, & -\frac{2\pi}{a} \le y < 0, \\ \int_{-\frac{2\pi}{a}+y}^{0} \frac{a^2}{8\pi^2}\left(1+\frac{a}{2\pi}z\right)dz + \int_{-\frac{\pi}{a}+y}^{\frac{\pi}{a}} \frac{a^2}{8\pi^2}\left(1-\frac{a}{2\pi}z\right)dz, & 0 \le y < \frac{2\pi}{a}, \\ \int_{-\frac{2\pi}{a}+y}^{\frac{2\pi}{a}} \frac{a^2}{8\pi^2}\left(1-\frac{a}{2\pi}z\right)dz, & \frac{2\pi}{a} \le y \le \frac{4\pi}{a}, \\ 0, & \textit{otherwise}, \end{cases} \tag{B.18}$$

which concludes the proof.

Lemma 11 *If the PDF of the sum of three independent random variables, $Y = 2\theta_l - (\theta_m + \theta_n)$, is given by* (B.16)*, then*

$$\mathbb{E}\left[\cos(2\theta_l - (\theta_m + \theta_n))\right] = \frac{a^3 \cos(\frac{\pi}{a}) \sin^3(\frac{\pi}{a})}{\pi^3}. \tag{B.19}$$

Proof 9 *By using Lemma 10 we have*

$$\begin{aligned} \mathbb{E}\left[\cos(2\theta_l - (\theta_m + \theta_n))\right] &= \mathbb{E}\left[\cos(y)\right] \\ &= \int_{-\frac{4\pi}{a}}^{-\frac{2\pi}{a}} \cos(y)\left[\frac{a}{2\pi} + \frac{a^2 y}{4\pi^2} + \frac{a^3 y^2}{32\pi^3}\right]dy \\ &\quad + \int_{-\frac{2\pi}{a}}^{\frac{2\pi}{a}} \cos(y)\left[\frac{a}{4\pi} - \frac{a^3 y^2}{32\pi^3}\right]dy \\ &\quad + \int_{\frac{2\pi}{a}}^{\frac{4\pi}{a}} \cos(y)\left[\frac{a}{2\pi} - \frac{a^2 y}{4\pi^2} + \frac{a^3 y^2}{32\pi^3}\right]dy. \end{aligned} \tag{B.20}$$

Solving the three integrals in (B.20) *concludes the proof.*

Lemma 12 *If θ_l, θ_m, and θ_n are independent and identically distributed uniform random variables with PDF given by* (B.7)*, then*

$$\mathbb{E}\left[\cos(\theta_l - \theta_m)\cos(\theta_l - \theta_n)\right] = \frac{a^2 \sin^2(\frac{\pi}{a})\left[2\pi + a\sin(\frac{2\pi}{a})\right]}{4\pi^3}. \tag{B.21}$$

Proof 10 *We start by applying the trigonometric identity $\cos(a)\cos(b) = \frac{\cos(a-b)+\cos(a+b)}{2}$ to* (B.22)*, which then can be re-written as*

$$\begin{aligned} \mathbb{E}\left[\cos(\theta_l - \theta_m)\cos(\theta_l - \theta_n)\right] &= \frac{1}{2}\mathbb{E}\left[\cos(\theta_n - \theta_m)\right] \\ &\quad + \frac{1}{2}\mathbb{E}\left[\cos(2\theta_l - \theta_n - \theta_m)\right]. \end{aligned} \tag{B.22}$$

Next, by applying Lemmas 7 and 11 to (B.22)*, we conclude the proof.*

B.1.1 Approximated PDF of the Instantaneous SNR

Let the random variable $Z = r$, where r is defined in (5.6), therefore, the PDF of Z can be accurately approximated by a Gamma distribution with parameters κ and θ, defined by (5.7) and (5.8), respectively. This is empirically proven by comparing the normalized histogram of Z against the theoretical PDF of a Gamma random variable, Y, with the parameters defined earlier.

In order to approximate Z as a Gamma random variable, Y, we have to find the parameters shape and scale (*i.e.*, κ and θ) based on statistical information of Z. Therefore, we approximate Z as a Gamma random variable, Y, by using two different moments of Y and then assuming that $\mathbb{E}[Y^2] = \mathbb{E}[Z^2]$ and $\mathbb{E}[Y^4] = \mathbb{E}[Z^4]$.

Those two moments of the Gamma distribution Y are defined as

$$\mathbb{E}\left[Y^2\right] = \kappa(\kappa+1)\theta^2, \tag{B.23}$$

and

$$\mathbb{E}\left[Y^4\right] = \kappa(\kappa+1)(\kappa+2)(\kappa+3)\theta^4. \tag{B.24}$$

Based on (B.23), the assumption that $\mathbb{E}[Y^2] = \mathbb{E}[Z^2]$ and then isolating θ we find

$$\theta = \sqrt{\frac{\mathbb{E}[Z^2]}{\kappa(\kappa+1)}}. \tag{B.25}$$

Next, plugging (B.25) back into (B.24) and assuming that $\mathbb{E}[Y^4] = \mathbb{E}[Z^4]$, we find κ as

$$\left(\mathbb{E}\left[Z^4\right] - \mathbb{E}\left[Z^2\right]^2\right)\kappa^2 + \left(\mathbb{E}\left[Z^4\right] - 5\mathbb{E}\left[Z^2\right]^2\right)\kappa - 6\mathbb{E}\left[Z^2\right]^2 = 0, \tag{B.26}$$

which is a quadratic equation with the following two roots

$$\kappa_0 = \frac{-\left(\mathbb{E}[Z^4] - 5\mathbb{E}[Z^2]^2\right) + \sqrt{\mathbb{E}[Z^4]^2 - 34\mathbb{E}[Z^4]\,\mathbb{E}[Z^2]^2 + 49\mathbb{E}[Z^2]^4}}{2\left(\mathbb{E}[Z^4] - \mathbb{E}[Z^2]^2\right)}, \tag{B.27}$$

$$\kappa_1 = \frac{-\left(\mathbb{E}[Z^4] - 5\mathbb{E}[Z^2]^2\right) - \sqrt{\mathbb{E}[Z^4]^2 - 34\mathbb{E}[Z^4]\,\mathbb{E}[Z^2]^2 + 49\mathbb{E}[Z^2]^4}}{2\left(\mathbb{E}[Z^4] - \mathbb{E}[Z^2]^2\right)}, \tag{B.28}$$

where out of the two roots, only is useful, *i.e.*, only one root has a positive value. Since κ ought to be a real and positive number, we assume that the value within the square root is a positive one. Next, assuming that $5\mathbb{E}[Z^2]^2 \geq \mathbb{E}[Z^4]$, then only κ_0 results in a positive value.

Next, in order to find the moment $\mathbb{E}[Z^2]$, we first expand $\mathbb{E}[Z^2]$ as

$$\begin{aligned}
\mathbb{E}\left[Z^2\right] &= \mathbb{E}[\gamma] \\
&= \mathbb{E}\left[\rho\left|\sum_{n=1}^{N}|h_n||g_n|e^{j\delta_n}\right|^2\right] \\
&= \mathbb{E}\left[\rho\sum_{n=1}^{N}d_n^2 + 2\rho\sum_{m=1}^{N}\sum_{n=m+1}^{N}d_m d_n\cos(\delta_m - \delta_n)\right],
\end{aligned} \tag{B.29}$$

where $d_k = |h_k||g_k|$ and the last line is found by applying Lemma 2. Thus, using the fact that $|h_n|$, $|g_n|$, and δ_n, $\forall n$ are mutually independent random variables and that h_m and h_n, and g_m and g_n, $\forall m, n$ are identically distributed, then (B.29), can be re-written as

$$\mathbb{E}\left[Z^2\right] = \rho \sum_{n=1}^{N} \mathbb{E}\left[|h_n|^2\right] \mathbb{E}\left[|g_n|^2\right] + 2\rho \sum_{m=1}^{N} \sum_{n=m+1}^{N} \mathbb{E}\left[|h_m|\right]^2 \mathbb{E}\left[|g_m|\right]^2 \mathbb{E}\left[\cos(\delta_m - \delta_n)\right]. \tag{B.30}$$

Then, by applying Lemmas 4 and 7 to (B.30), we find (5.11).

Next, in order to find the moment $\mathbb{E}\left[Z^4\right]$, we initially expand it as

$$\begin{aligned} \mathbb{E}\left[Z^4\right] &= \mathbb{E}\left[\gamma^2\right] \\ &= \mathbb{E}\left[\left(\sum_{l=1}^{N} d_l^2\right)^2\right] + 4\sum_{l=1}^{N}\sum_{m=1}^{N}\sum_{n=m+1}^{N} \mathbb{E}\left[d_l^2 d_m d_n \cos(\delta_m - \delta_n)\right] \\ &\quad + 4\mathbb{E}\left[\left(\sum_{m=1}^{N}\sum_{n=m+1}^{N} d_m d_n \cos(\delta_m - \delta_n)\right)^2\right], \end{aligned} \tag{B.31}$$

where $d_k = |h_k||g_k|$. The first term of (B.31) can be expressed as

$$\begin{aligned} \mathbb{E}\left[\left(\sum_{l=1}^{N} d_l^2\right)^2\right] &= \mathbb{E}\left[\sum_{n=1}^{N} d_n^4 + \sum_{m=1}^{N}\sum_{n=1,n\neq m}^{N} d_m^2 d_n^2\right] N\mathbb{E}\left[|g_m|^4\right] \mathbb{E}\left[|h_m|^4\right] \\ &\quad + N \quad (N-1)\mathbb{E}\left[|g_m|^2\right]^2 \mathbb{E}\left[|h_m|^2\right]^2 \\ &= N(N+3)\left(\beta_g \beta_h\right)^2, \end{aligned} \tag{B.32}$$

where the last line of (B.32) is found by applying Lemma 4. Next, the second term of (B.31) can be expressed as (B.33), where the last line is found by applying Lemmas 4 and 7. Then, the third term of (B.31) can be expressed as (B.34), where the last line is found after applying Lemmas 4, 7, 8, and 12. Finally, after plugging (B.32), (B.33), and (B.34) back into (B.31) and several simplifications, we find (5.12).

The proof is concluded by replacing Equations (5.11) and (5.12) into the definitions of κ and θ, given by (B.27) and (B.25), respectively.

B.2 Derivation of Remark 1

For the derivation of Remark 1, we need to define the following Lemma.

Lemma 13

$$\lim_{x\to 0} \frac{\sin(x)}{x} = 1. \tag{B.35}$$

Proof 11 *We prove Lemma 13 by applying L'Hôpital's rule to* (B.35) *as shown next*

$$\lim_{x\to 0} \frac{\frac{\partial \sin(x)}{\partial x}}{\frac{\partial x}{\partial x}} = \lim_{x\to 0} \cos(x) = 1. \tag{B.36}$$

$$
\begin{aligned}
4\sum_{l=1}^{N}\sum_{m=1}^{N}\sum_{n=m+1}^{N}\mathbb{E}\left[d_l^2 d_m d_n\cos(\delta_m-\delta_n)\right] &= 4\sum_{m=1}^{N}\sum_{n=1,n\neq m}^{N}\mathbb{E}\left[d_m^3 d_n\cos(\delta_m-\delta_n)\right] \\
&\quad + 4\sum_{l=1}^{N}\sum_{m=1,m\neq l}^{N}\sum_{n=m+1,n\neq l}^{N}\mathbb{E}\left[d_l^2 d_m d_n\cos(\delta_m-\delta_n)\right] \\
&= 4N(N-1)\mathbb{E}\left[|g_m|^3\right]\mathbb{E}\left[|h_m|^3\right]\mathbb{E}\left[|g_m|\right]\mathbb{E}\left[|h_m|\right]\mathbb{E}\left[\cos(\delta_l-\delta_m)\right] \\
&\quad + 2N(N-1)(N-2)\mathbb{E}\left[|g_l|^2\right]\mathbb{E}\left[|h_l|^2\right]\mathbb{E}\left[|g_m|\right]\mathbb{E}\left[|h_m|\right]\mathbb{E}\left[|g_n|\right]\mathbb{E}\left[|h_n|\right]\cos(\delta_m-\delta_n) \\
&= \frac{1}{16}N(N-1)(2N+5)\left(\beta_g\beta_h\right)^2 Q^2\sin^2\left(\frac{\pi}{Q}\right).
\end{aligned}
\tag{B.33}
$$

$$
\begin{aligned}
4\mathbb{E}\left[\left(\sum_{m=1}^{N}\sum_{n=m+1}^{N}d_m d_n\cos(\delta_m-\delta_n)\right)^2\right] &= 4\sum_{j=1}^{N}\sum_{l=j+1}^{N}\sum_{m=1}^{N}\sum_{n=m+1}^{N}\mathbb{E}\left[d_j d_l d_m d_n\cos(\delta_j-\delta_l)\cos(\delta_m-\delta_n)\right] \\
&= 4\sum_{j=1}^{N}\sum_{l=j+1}^{N}\sum_{m=1,m=j}^{N}\sum_{n=m+1,n=l}^{N}\mathbb{E}\left[d_j^2 d_l^2\cos^2(\delta_j-\delta_l)\right] \\
&\quad + 8\sum_{l=1}^{N}\sum_{m=1,m\neq l}^{N}\sum_{n=m+1,n\neq l}^{N}\mathbb{E}\left[d_l^2 d_m d_n\cos(\delta_l-\delta_m)\cos(\delta_l-\delta_n)\right] \\
&\quad + 4\sum_{j=1}^{N}\sum_{l=1,l\neq j\neq m\neq n}^{N}\sum_{m=1,l\neq j\neq m\neq n}^{N}\sum_{n=1,l\neq j\neq m\neq n}^{N}\mathbb{E}\left[d_j d_l d_m d_n\cos(\delta_j-\delta_l)\cos(\delta_m-\delta_n)\right] \\
&= \frac{N(N-1)\left(\beta_g\beta_h\right)^2\left\{Q^2\left[\pi(N-2)\sin^2\left(\frac{\pi}{Q}\right)\left(\pi\left((N-3)Q^2\sin^2\left(\frac{\pi}{Q}\right)+32\right)+16Q\sin\left(\frac{2\pi}{Q}\right)\right)-32\cos\left(\frac{4\pi}{Q}\right)\right]+32\left(Q^2+8\pi^2\right)\right\}}{256\pi^2}.
\end{aligned}
\tag{B.34}
$$

Lemma 14

$$
\lim_{x\to\infty}\left(x\sin\left(\frac{a}{x}\right)\right)^n = a^n, \forall a, n\in\mathbb{R}. \tag{B.37}
$$

Proof 12 *We start by re-writing* (B.37) *as*

$$
\left(\lim_{x\to\infty}\frac{a}{a}\frac{\sin\left(\frac{a}{x}\right)}{\frac{1}{x}}\right)^n = \left(\lim_{x\to\infty}a\frac{\sin\left(\frac{a}{x}\right)}{\frac{a}{x}}\right)^n, \tag{B.38}
$$

where we also used the power rule of limits to re-write it. Next, we apply the following change of variables $\theta=\frac{a}{x}$ *to* (B.38)*, resulting in*

$$
\left(\lim_{\theta\to 0}a\frac{\sin(\theta)}{\theta}\right)^n = a^n\left(\lim_{\theta\to 0}\frac{\sin(\theta)}{\theta}\right)^n, \tag{B.39}
$$

where we used the constant multiple rule of limits to find the last part. Next, by using Lemma 13, we conclude the proof.

B.2.1 Derivation

The results (5.13) and (5.14) are found after expanding (5.11) and (5.12), using Lemma 14 and the fact that $\lim_{x\to\infty}\cos(1/x)=1$.

B.3 Proof of (5.17)

For the proof of (5.17), we should notice that when $\lim_{\rho\to\infty}\theta = \infty$ then, consequently, $\lim_{\rho\to\infty} -\frac{1}{4\theta^2} = 0$. Therefore,

$$\lim_{\rho\to\infty} {}_2F_3\left(1,1;2,\frac{3}{2}-\frac{\kappa}{2},2-\frac{\kappa}{2};-\frac{1}{4\theta^2}\right) = 1. \tag{B.40}$$

$$\lim_{\rho\to\infty} {}_1F_2\left(\frac{\kappa}{2}+\frac{1}{2};\frac{3}{2},\frac{\kappa}{2}+\frac{3}{2};-\frac{1}{4\theta^2}\right) = 1. \tag{B.41}$$

$$\lim_{\rho\to\infty} {}_1F_2\left(\frac{\kappa}{2};\frac{1}{2},\frac{\kappa}{2}+1;-\frac{1}{4\theta^2}\right) = 1. \tag{B.42}$$

Hence, in high SNR regime (5.16) can be tightly approximated as (5.17), which concludes the proof.

B.4 Proof of (5.18)

The proof of (5.18) is straightforwardly found by noticing that the first three terms of (5.17) tend to 0 when $\rho \to \infty$, since $\theta \to \infty$ when $\rho \to \infty$, which concludes this proof.

B.5 High SNR regime

In high SNR regime, as $N \to \infty$ and $\kappa \to \infty$, $\Gamma(\kappa)$ grows even faster. Therefore,

$$\lim_{N\to\infty} \frac{2\pi \sec\left(\frac{\pi\kappa}{2}\right)}{(\kappa+1)\theta^{\kappa+1}\Gamma(\kappa)\log(4)} = 0. \tag{B.43}$$

$$\lim_{N\to\infty} \frac{2\pi \csc\left(\frac{\pi\kappa}{2}\right)}{\kappa\theta^{\kappa}\Gamma(\kappa)\log(4)} = 0. \tag{B.44}$$

These two terms tend to 0 faster than the other 2 terms, concluding the proof.

B.6 Derivation of C_{lower}

Here we outline the derivation of C_{lower} in (5.25). We start by applying the Taylor series expansion of $1/\gamma$ around $\mathbb{E}[\gamma]$ [70], the term $\mathbb{E}[1/\gamma]$ in (5.24) can be approximated as [79]

$$\mathbb{E}\left[\frac{1}{\gamma}\right] \approx \frac{1}{\mathbb{E}[\gamma]} + \frac{\text{var}(\gamma)}{\mathbb{E}[\gamma]^3} = \frac{\mathbb{E}[\gamma^2]}{\mathbb{E}[\gamma]^3}. \tag{B.45}$$

After replacing $\mathbb{E}[\gamma]$ and $\mathbb{E}[\gamma^2]$ in (B.45) via (5.11) and (5.12), respectively, and then by substituting the resultant expression into (5.24), C_{lower} can be approximated as shown in the second part of (5.25).

B.7 Derivation of the outage probability

Here we describe the derivation of the outage probability given by (5.34). Using the PDF of the instantaneous capacity given by (5.20), the outage probability can be written as

$$\begin{aligned} P_{\text{out.}} &= \Pr\{C_{\text{inst.}} < C_{\text{out.}}\} \\ &= \frac{\log(2)}{\Gamma(\kappa)\theta^{\kappa}} \int_0^{C_{\text{out.}}} 2^{u-1}(2^u-1)^{\left(\frac{\kappa-2}{2}\right)} e^{-\frac{\sqrt{2^u-1}}{\theta}} du. \end{aligned} \tag{B.46}$$

Next, using the following change of variable $x = 2^u - 1$, then (B.46) becomes

$$P_{\text{out.}} = \frac{1}{2\Gamma(\kappa)\theta^{\kappa}} \int_0^{2^{C_{\text{out.}}}-1} x^{\left(\frac{\kappa-2}{2}\right)} e^{-\frac{\sqrt{x}}{\theta}} dx. \tag{B.47}$$

Finally, using (2.33.10) from [70]

$$\int x^m e^{-\beta x^n} dx = -\frac{\Gamma\left(\frac{m+1}{n}, \beta x^n\right)}{n\beta^{\frac{m+1}{n}}}, \tag{B.48}$$

we find a solution for the the integral in (B.47), which concludes the proof.

B.8 Proofs of (5.36) and (5.37)

For the proofs of (5.36) and (5.37), we first to define the following Lemmas.

Lemma 15 *According to (Eq. 8.2.5, [71])*

$$1 - \frac{\Gamma(a,b)}{\Gamma(a)} = \frac{\gamma(a,b)}{\Gamma(a)}, \tag{B.49}$$

where $\gamma(a,b)$ is the lower incomplete gamma function.

Lemma 16 *According to (Eq. 8.5.1, [71])*

$$\gamma(a,b) = a^{-1}b^a\, {}_1F_1(a, a+1, -b). \tag{B.50}$$

Lemma 17 *According to (07.20.03.0001.01) of [80]*

$${}_1F_1(a,b,0) = 1. \tag{B.51}$$

Therefore, applying Lemmas 15 and 16, defined above, to (5.35) we end up with (5.36), which concludes the proof. Now, (5.37) is found by applying Lemma 17 to (5.36) and remembering that $\lim_{\rho\to\infty}\theta = \infty$, then $\lim_{\rho\to\infty} 1/\theta = 0$.

B.9 Derivation of (5.39)

In this Appendix, we derive the average symbol error rate expression given by (5.39), but first, we need to establish some Lemmas.

Lemma 18

$$\mathcal{Q}(x) = \frac{1}{2}\left[1 - erf\left(\frac{x}{\sqrt{2}}\right)\right]. \tag{B.52}$$

This relation is given by (Eq. B.111, [81]).

Lemma 19

$$\int_0^\infty x^m e^{-\beta x^n} dx = \frac{\Gamma\left(\frac{m+1}{n}\right)}{n\beta^{\frac{m+1}{n}}}. \tag{B.53}$$

This relation is given by (Eq. 3.326.2, [70]).

Lemma 20 *If erf(.) is the Gauss error function, and a, b, and $c > 0$, then the integral $\int_0^\infty erf(ax)\, x^b e^{-cx} dx$ is given by* (B.54). *The integral in* (B.54) *is found by using an integral solver [82].*

$$\begin{aligned}\int_0^\infty erf(ax)\, x^b e^{-cx} dx &= c^{-b-1}\Gamma(b+1) \\ &+ \frac{ca^{-b-2}\Gamma\left(\frac{b+3}{2}\right)\, {}_2F_2\left(\frac{b}{2}+1, \frac{b}{2}+\frac{3}{2}; \frac{3}{2}, \frac{b}{2}+2; \frac{c^2}{4a^2}\right)}{\sqrt{\pi}(b+2)} \\ &- \frac{a^{-b-1}\Gamma\left(\frac{b}{2}+1\right)\, {}_2F_2\left(\frac{b}{2}+\frac{1}{2}, \frac{b}{2}+1; \frac{1}{2}, \frac{b}{2}+\frac{3}{2}; \frac{c^2}{4a^2}\right)}{\sqrt{\pi}(b+1)}.\end{aligned} \tag{B.54}$$

B.9.1 Proof of the Average Symbol Error Rate

By using the fact that $\gamma = r^2$ (see (5.6)), the expectation of the conditional symbol error probability given the distribution of the SNR can be written as

$$\begin{aligned}P_e &= \mathbb{E}\left[a\mathcal{Q}\left(\sqrt{b\gamma}\right)\right] = \mathbb{E}\left[a\mathcal{Q}\left(\sqrt{b}r\right)\right] \\ &= \int_0^\infty P_{e|\gamma}(x) f_R(x) dx,\end{aligned} \tag{B.55}$$

where $f_R(r)$ is the PDF of the Gamma distribution, which tightly approximates the exact PDF of the random variable, r.

By plugging $P_{e|\gamma} = a\mathcal{Q}\left(\sqrt{b\gamma}\right)$ and the Gamma PDF back into (B.55), the average SER is rewritten as

$$P_e = \frac{a}{\Gamma(\kappa)\theta^\kappa} \int_0^\infty \mathcal{Q}\left(\sqrt{b}x\right) x^{\kappa-1} e^{-x/\theta} dx. \tag{B.56}$$

By using Lemma 18, (B.56) can be equivalently rewritten as

$$P_e = \frac{a}{\Gamma(\kappa)\theta^\kappa}\left[\int_0^\infty x^{\kappa-1}e^{-x/\theta}dx - \int_0^\infty \operatorname{erf}\left(\sqrt{\frac{b}{2}}x\right)x^{\kappa-1}e^{-x/\theta}dx\right]. \tag{B.57}$$

The first integral inside the square brackets of (B.57) is found by applying Lemma 19 to it, which results in

$$\int_0^\infty x^{\kappa-1}e^{-x/\theta}dx = \Gamma(\kappa)\theta^\kappa. \tag{B.58}$$

The second integral inside the square brackets of (B.57) is found by applying Lemma 20 to it, which results in

$$\begin{aligned}&\int_0^\infty \operatorname{erf}\left(\sqrt{\frac{b}{2}}x\right)x^{\kappa-1}e^{-x/\theta}dx\\ &= \theta^\kappa\Gamma(\kappa) - \frac{2^{\kappa/2}b^{-\frac{\kappa}{2}}\Gamma\left(\frac{\kappa+1}{2}\right)\,{}_2F_2\left(\frac{\kappa}{2}+\frac{1}{2},\frac{\kappa}{2};\frac{1}{2},\frac{\kappa}{2}+1;\frac{1}{2b\theta^2}\right)}{\sqrt{\pi}\kappa}\\ &+ \frac{2^{\frac{\kappa}{2}+\frac{1}{2}}b^{-\frac{\kappa}{2}-\frac{1}{2}}\Gamma\left(\frac{\kappa}{2}+1\right)\,{}_2F_2\left(\frac{\kappa}{2}+\frac{1}{2},\frac{\kappa}{2}+1;\frac{3}{2},\frac{\kappa}{2}+\frac{3}{2};\frac{1}{2b\theta^2}\right)}{\sqrt{\pi}\theta(\kappa+1)}.\end{aligned} \tag{B.59}$$

Finally, by substituting (B.58) and (B.59) back into (B.57), we conclude the proof.

B.10 Proof of (5.40)

For the proof of (5.40) we should notice that when $\lim_{\rho\to\infty}\theta = \infty$ then, consequently, $\lim_{\rho\to\infty}\frac{1}{2\theta^2} = 0$. Therefore,

$$\lim_{\rho\to\infty}\,{}_2F_2\left(\frac{\kappa}{2}+\frac{1}{2},\frac{\kappa}{2};\frac{1}{2},\frac{\kappa}{2}+1;\frac{1}{2b\theta^2}\right) = 1. \tag{B.60}$$

$$\lim_{\rho\to\infty}\,{}_2F_2\left(\frac{\kappa}{2}+\frac{1}{2},\frac{\kappa}{2}+1;\frac{3}{2},\frac{\kappa}{2}+\frac{3}{2};\frac{1}{2b\theta^2}\right) = 1. \tag{B.61}$$

Hence, in high SNR regime (5.39) can be tightly approximated as (5.40), which concludes the proof.

B.11 Derivation of diversity order

In order to derive the diversity order, we first need to rewrite (5.11) and (5.12) as

$$\mathbb{E}\left[\gamma\right] = \rho\beta_g\beta_h\mathcal{A}_1, \tag{B.62}$$

and

$$\mathbb{E}\left[\gamma^2\right] = \left(\rho\beta_g\beta_h\right)^2 \mathcal{A}_2, \tag{B.63}$$

respectively, where $\mathcal{A}_1$ and $\mathcal{A}_2$ do not depend on the average SNR, ρ. By plugging these two equation back into (5.7) and (5.8), we find

$$\kappa = \frac{5\mathcal{A}_1^2 + \sqrt{49\mathcal{A}_1^4 - 34\mathcal{A}_1^2\mathcal{A}_2 + \mathcal{A}_2^2} - \mathcal{A}_2}{2\left(\mathcal{A}_2 - \mathcal{A}_1^2\right)} > 0, \tag{B.64}$$

which also does not depend on the average SNR, and

$$\begin{aligned}\theta &= \rho^{\frac{1}{2}}\sqrt{\frac{\beta_g\beta_h\left(2\mathcal{A}_1^2 - \sqrt{\mathcal{A}_1^4 + 14\mathcal{A}_1^2\mathcal{A}_2 + \mathcal{A}_2^2} + 2\mathcal{A}_2\right)}{6\mathcal{A}_1}} \\ &= \rho^{\frac{1}{2}}\theta' > 0,\end{aligned} \tag{B.65}$$

which depends on the average SNR. Therefore, in high-SNR regime, (5.40) can be written as

$$P_e^{\text{high-SNR}} \approx \mathcal{B}_1\rho^{-\frac{\kappa}{2}} - \mathcal{B}_2\rho^{-\frac{(\kappa+1)}{2}}, \tag{B.66}$$

where

$$\mathcal{B}_1 = \frac{a2^{-\frac{(\kappa+2)}{2}}b^{-\frac{\kappa}{2}}\theta'^{-\kappa}}{\Gamma\left(\frac{\kappa}{2}+1\right)}, \tag{B.67}$$

and

$$\mathcal{B}_2 = \frac{\kappa a2^{-\frac{(\kappa+3)}{2}}b^{-\frac{(\kappa+1)}{2}}\theta'^{-(\kappa+1)}}{\Gamma\left(\frac{\kappa+3}{2}\right)}. \tag{B.68}$$

Note that $\mathcal{B}_1$ and $\mathcal{B}_2$ do not depend on the average SNR, *i.e.*, they are independent from it. Furthermore, from (B.66), we realise that the terms $\rho^{-\frac{\kappa}{2}}$ and $\rho^{-\frac{(\kappa+1)}{2}}$ contribute with diversity order of $\frac{\kappa}{2}$ and $\frac{(\kappa+1)}{2}$, respectively. Therefore, the diversity order is calculated as

$$D = \min\left(\frac{\kappa}{2}, \frac{(\kappa+1)}{2}\right). \tag{B.69}$$

Since $\kappa > 0$, then (B.69) is simplified as

$$D = \frac{\kappa}{2}. \tag{B.70}$$

The proof is concluded after plugging (B.64) into (B.70).

List of Publications

During my Ph.D., we have published the following works:

- M. S. P. Facina and G. Fraidenraich, Accurate log-normal approximation to the signal-to-interference ratio in massive multiple-input multiple-output systems, *Internet Technology Letters*, 2018, 1:e48.
- M. S. P. Facina and G. Fraidenraich "Preliminary Studies on the Large Intelligent Surfaces Efficiencies Under Different Channels," *XXXVIII Brazilian Symposium on Telecommunications and Signal Processing*, 2020.
- R. C. Ferreira, M. S. P. Facina, F. A. P. de Figueiredo and G. Fraidenraich, and E. R. de Lima, "Bit Error Probability for Large Intelligent Surfaces Under Double-Nakagami Fading Channels," *IEEE Open Journal of the Communications Society*, 2020.
- F. A. P. de Figueiredo, M. S. P. Facina, R. C. Ferreira and G. Fraidenraich, "Large Intelligent Surfaces With Discrete Set of Phase-Shifts Communicating Through Double-Rayleigh Fading Channels", *IEEE Access*, submitted.
- R. C. Ferreira, M. S. P. Facina, F. A. P. de Figueiredo and G. Fraidenraich, "Large Intelligent Surfaces Communicating Through Massive MIMO Rayleigh Fading Channels has equivalent Gamma fading", *IEEE Sendors Journal*, accepted.